

(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

CARL MURCHISON

GORDON ALLPORT
Harvard University

ARTHUR G. BILLS
University of Cincinnati

E. A. BOTT
University of Toronto

LEONARD CARMICHAEL
Smithsonian Institution

HERBERT S. CONRAD
U. S. Office of Education

ELMER K. CULLER
University of Rochester

JOHN F. DASHIELL
University of North
Carolina

ROY M. DORCUS
University of California
at Los Angeles

HORACE B. ENGLISH
Ohio State University

FRANKLIN FEARING
University of California
at Los Angeles

If this space should be unstamped, this is the regular library edition.
But if this space is stamped with a designating title, this is a special

edition, sold under the restrictions of a bilateral contract, and may not
be resold for a period of five years from date of publication.

FRANK A. GELDARD
University of Virginia

CLARENCE H. GRAHAM
Columbia University

ERNEST R. HILGARD
Stanford University

HEINRICH KLÜVER
University of Chicago

CARNEY LANDIS
Columbia University

NORMAN L. MUNN
University of Adelaide,
Australia

JOHN PAUL NAFE
Florida State University

DONALD G. PATERSON
University of Minnesota

LEE EDWARD TRAVIS
University of Southern
California

ROBERT C. TRYON
University of California

MICHAEL J. ZIGLER
Wellesley College

VOLUME 56

1963

Copyright, 1963, by The Journal Press

All rights reserved

Published Quarterly by The Journal Press
Provincetown, Massachusetts, U. S. A.

AUTHOR INDEX

| | | | |
|------------------------------|----------------------------|-------------------------------|-------------------|
| Abramson, H. A. | 363 | Knapp, Robert H. | 473 |
| Alexander, Theron | 273 | Knobloch, Hilda | 73 |
| Allison, Roger B., Jr. | 115 | Krampen, Martin | 61 |
| Averill, James R. | 455 | Krasner, Leonard | 193 |
| Baird, J. C. | 165 | Lachman, Sheldon J. | 239 |
| Bartley, S. Howard | 19, 99, 107, 175, 219, 435 | Lana, Robert E. | 9 |
| Becknell, James C., Jr. | 165 | Lawson, Reed | 203 |
| Bourassa, Charles M. | 435 | Lepley, William M. | 69 |
| Brown, Donald W. | 413 | Manegold, Richard F. | 357 |
| Caldwell, Bettye M. | 375 | Mayzner, M. S. | 161 |
| Caldwell, Fred T., Jr. | 375 | McGee, Mother Jane Marie | 447 |
| Cowen, Emory L. | 307 | McPherson, D. Angus | 137 |
| Davis, Roger T. | 119 | Miles, Harold | 307 |
| Day, David R. | 3 | Moldovan, Stanley E. | 375 |
| Dean, G. | 363 | Mulholland, Thomas | 227, 233 |
| Dietze, Alfred G. | 143 | Nelson, Thomas M. | 99, 107, 175, 219 |
| DiMascio, Alberto | 89 | Noer, David | 15 |
| Ehlinger, Linda A. | 473 | Pasamanick, Benjamin | 73 |
| Farnsworth, Paul R. | 269 | Pierson, George R. | 185, 441 |
| Fitzgerald, Maureen P. | 405 | Porter, Paul B. | 79 |
| Flescher, Irwin | 251 | Priddy, Ronald D. | 455 |
| Gabriel, R. F. | 161 | Ranney, Jane E. | 19 |
| Gengerelli, J. A. | 455 | Robertson, Malcolm H. | 213 |
| Gettner, H. H. | 363 | Rolo, A. | 363 |
| Gilbert, Albin R. | 385 | Scodel, Alvin | 429 |
| Goode, Omar S. | 3 | Shapiro, Stewart B. | 479 |
| Gregor, A. James | 137 | Steele, James P. | 119 |
| Gustav, Alice | 421 | Stogdill, Ralph M. | 3 |
| Izzo, Louis D. | 307 | Telschow, Earl F. | 307 |
| Jencks, Beata | 79 | Trost, Mary Ann | 307 |
| Jennings, Joseph W. | 43, 53 | Ullmann, Leonard | 193 |
| Jewell, Robert M. | 99, 107 | Weinstein, Will K. | 203 |
| Johnson, Donald M. | 43, 53 | Whittaker, James O. | 15 |
| Kassarjian, Harold H. | 85 | Wilson, Warner R. | 165 |
| Kelly, Robert F. | 185, 441 | Wise, Roberta F. | 219 |
| Kidd, Aline H. | 29 | Wolter, Douglas J. | 213 |
| King, David J. | 357 | Yavuz, Halide S. | 171 |
| King, H. E. | 283, 291, 299 | Zax, Melvin | 307 |

TABLE OF CONTENTS

| | |
|--|-----|
| The leader behavior of United States Senators | 3 |
| BY RALPH M. STOGDILL, OMAR S. GOODE, AND DAVID R. DAY | |
| Interest, media, and order effects in persuasive communications | 9 |
| BY ROBERT E. LANA | |
| Effects of masculine-feminine ego-involvements on the acquisition of a mirror-tracing skill | 15 |
| BY DAVID NOER AND JAMES O. WHITTAKER | |
| A further study of determinants of phenomenal distance in plane targets perceived as three-dimensional scenes | 19 |
| BY JANE E. RANNEY AND S. HOWARD BARTLEY | |
| The effect of stimulus color and content upon schizophrenics | 29 |
| BY ALINE H. KIDD | |
| Serial analysis of three problem-solving processes | 43 |
| BY D. M. JOHNSON AND J. W. JENNINGS | |
| Context effects in production and judgment | 53 |
| BY JOSEPH W. JENNINGS AND DONALD M. JOHNSON | |
| Handedness as a variable of importance in determining apparent movement direction | 61 |
| BY MARTIN KRAMPEN | |
| "The maturity of the chances": A gambler's fallacy | 69 |
| BY WILLIAM M. LEPLEY | |
| Early feeding and birth difficulties in childhood schizophrenia: An explanatory note | 73 |
| BY BENJAMIN PASAMANICK AND HILDA KNOBLOCH | |
| Need reduction and reinforcement: Incentive values of various goal objects and novelty | 79 |
| BY BEATA JENCKS AND PAUL B. PORTER | |
| Voting intentions and political perception | 85 |
| BY HAROLD H. KASSARJIAN | |
| Drug effects on competitive-paired associate learning: Relationship to and implications for the Taylor Manifest Anxiety Scale | 89 |
| BY ALBERTO DiMASCIO | |
| Effects upon brightness produced by varying the length of the null interval separating successive "single" pulses: Sensory implications of the Alternation of Response Theory, I | 99 |
| BY THOMAS M. NELSON, S. HOWARD BARTLEY, AND ROBERT M. JEWELL | |
| Brightness changes associated with variation in the number of pulses in a repetitive photic train: Sensory implications of the Alternation of Response Theory, II | 107 |
| BY THOMAS M. NELSON, S. HOWARD BARTLEY, AND R. M. JEWELL | |
| Using adverbs as multipliers in semantic differentials | 115 |
| BY ROGER B. ALLISON, JR. | |

| | |
|--|-----|
| Performance selection through radiation death in rhesus monkeys | 119 |
| BY ROGER T. DAVIS AND JAMES P. STEELE | |
| The correlation of the Porteus Maze and the Gestalt Continuation as personnel selection tests of peripheral peoples | 137 |
| BY A. JAMES GREGOR AND D. ANGUS MCPHERSON | |
| Types of emotions or dimensions of emotion? A comparison of typal analysis with factor analysis | 143 |
| BY ALFRED G. DIETZE | |
| Information "chunking" and short-term retention | 161 |
| BY M. S. MAYZNER AND R. F. GABRIEL | |
| The effect of frequency of presentation on the choice of nonsense syllables . | 165 |
| BY JAMES C. BECKNELL, JR., WARNER R. WILSON, AND J. C. BAIRD | |
| The production of random letters sequences in schizophrenics | 171 |
| BY HALIDE S. YAVUZ | |
| Brightness from repetitive photic trains varying in pulse number and separation intervals: Sensory implications of the Alternation of Response Theory, III | 175 |
| BY THOMAS M. NELSON, S. HOWARD BARTLEY, AND ROBERT M. JEWELL | |
| HSPQ norms on a state-wide delinquent population | 185 |
| BY GEORGE R. PIERSON AND ROBERT F. KELLY | |
| Variables affecting report of awareness in verbal conditioning | 193 |
| BY LEONARD KRASNER AND LEONARD P. ULLMANN | |
| The effect of experimentally-induced "awareness" upon performance in free- operant verbal conditioning and on subsequent tests of "awareness" . . . | 203 |
| BY WILL K. WEINSTEIN AND REED LAWSON | |
| The effect of sensory deprivation upon scores on the Wechsler Adult Intel- ligence Scale | 213 |
| BY MALCOLM H. ROBERTSON AND DOUGLAS J. WOLTER | |
| Size discrimination under two conditions of photic intermittency | 219 |
| BY THOMAS M. NELSON, S. HOWARD BARTLEY, AND ROBERTA F. WISE | |
| Notes concerning "nonmotion" responses while viewing rotating stimulus objects | 227 |
| BY THOMAS MULHOLLAND | |
| Instructional sets and motion perceived while viewing rotating stimulus objects | 233 |
| BY THOMAS MULHOLLAND | |
| A behavioristic rationale for the development of psychosomatic phenomena . | 239 |
| BY SHELDON J. LACHMAN | |
| ✓ Anxiety and achievement of intellectually gifted and creatively gifted children . | 251 |
| BY IRWIN FLESCHER | |
| Has the status of music changed in 30 years? | 269 |
| BY PAUL R. FARNSWORTH | |
| The effect of psychopathology in children's drawing of the human figure . . | 273 |
| BY THERON ALEXANDER | |

| | |
|---|-----|
| The retention of sensory experience: I. Intensity | 283 |
| By H. E. KING | |
| The retention of sensory experience: II. Frequency | 291 |
| By H. E. KING | |
| The retention of sensory experience: III. Duration | 299 |
| By H. E. KING | |
| A preventive mental health program in the school setting: Description and evaluations | 307 |
| By EMORY L. COWEN, LOUIS D. IZZO, HAROLD MILES, EARL F. TELSCHOW, MARY ANN TROST, AND MELVIN ZAX | |
| Clinical judgment | 357 |
| By DAVID J. KING AND RICHARD F. MANEGOLD | |
| Lysergic acid diethylamide (LSD-25): XXXIV. Comparison with effect of psilocybin on the Siamese fighting fish | 363 |
| By H. A. ABRAMSON, H. H. GETTNER, A. ROLO, AND G. DEAN | |
| Emotional behavior of the white rat following full-thickness thermal burns | 375 |
| By FRED T. CALDWELL, JR., BETTYE M. CALDWELL, AND STANLEY E. MOLDOVAN | |
| Toward an automated technique of probing into emotional blocks | 385 |
| By ALBIN R. GILBERT | |
| Self-disclosure and expressed self-esteem, social distance and areas of the self revealed | 405 |
| By MAUREEN P. FITZGERALD | |
| The effect of observer redundancy on display monitoring equipment | 413 |
| By DONALD W. BROWN | |
| Response set in objective achievement tests | 421 |
| By ALICE GUSTAV | |
| Probability preferences and expected values | 429 |
| By ALVIN SCODEL | |
| A study of the possible effect of differential light adaptation on measures of brightness enhancement | 435 |
| By CHARLES M. BOURASSA AND S. HOWARD BARTLEY | |
| Anxiety, extraversion, and personality idiosyncrasy in delinquency | 441 |
| By GEORGE R. PIERSON AND ROBERT F. KELLY | |
| The effect of group verbal suggestion and age on the perception of the Ames trapezoidal illusion | 447 |
| By MOTHER JANE MARIE MCGEE | |
| Studies in the neurophysiology of learning: VII. Bar-pressing rate for cranial self-stimulation as related to frequency of rectangular pulses when Q is held constant | 455 |
| By J. A. GINGERELLI, RONALD D. PRIDDY, AND JAMES R. AYERILL | |
| Sex differences among schizophrenics in the interpretation of the human diad | 473 |
| By ROBERT H. KNAPP AND LINDA A. EHLINGER | |
| Transactional aspects of ego therapy | 479 |
| By STEWART B. SHAPIRO | |

\$20.00 per annum
\$15.00 per volume
\$7.50 per half volume

QUARTERLY
Two volumes per year
Immediate publication

July, 1963
Volume 56, First Half

(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition.* But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a liberal contract, and may not be resold for a period of five years from the date of publication.

JULY, 1963

(Manuscripts are printed in the order of final acceptance)

| | |
|---|----|
| The leader behavior of United States Senators | 3 |
| By RALPH M. STODDILL, OMAR S. GOODE, AND DAVID R. DAY | |
| Interest, media, and order effects in persuasive communications | 9 |
| By ROBERT E. LANA | |
| Effects of masculine-feminine ego-involvements on the acquisition of a mirror-tracing skill | 15 |
| By DAVID NOER AND JAMES O. WHITTAKER | |
| A further study of determinants of phenomenal distance in plane targets perceived as three-dimensional scenes | 19 |
| By JANE E. RANNEY AND S. HOWARD BARTLEY | |
| The effect of stimulus color and content upon schizophrenics | 29 |
| By ALINE H. KIDD | |
| Serial analysis of three problem-solving processes | 43 |
| By D. M. JOHNSON AND J. W. JENNINGS | |
| Context effects in production and judgment | 53 |
| By JOSEPH W. JENNINGS AND DONALD M. JOHNSON | |
| Handedness as a variable of importance in determining apparent movement direction | 61 |
| By MARTIN KRAMPEN | |

(OVER)

Copyright, 1963, by The Journal Press
Provincetown, Massachusetts, U. S. A.

Entered as second-class matter January 26, 1937, at the post-office at
Provincetown, Mass., under the Act of March 3, 1879
Second-class postage paid at Provincetown, Mass.

| | |
|--|-----|
| "The maturity of the chances": A gambler's fallacy | 69 |
| BY WILLIAM M. LEPLY | |
| Early feeding and birth difficulties in childhood schizophrenia: An explanatory note | 73 |
| BY BENJAMIN PASAMANICK AND HILDA KNOBLOCH | |
| Need reduction and reinforcement: Incentive values of various goal objects and novelty | 79 |
| BY BEATA JENCKS AND PAUL B. PORTER | |
| Voting intentions and political perception | 85 |
| BY HAROLD H. KASSARJIAN | |
| Drug effects on competitive-paired associate learning: Relationship to and implications for the Taylor Manifest Anxiety Scale | 89 |
| BY ALBERTO DIMASCIO | |
| Effects upon brightness produced by varying the length of the null interval separating successive "single" pulses: Sensory implications of the Alternation of Response Theory, I | 99 |
| BY THOMAS M. NELSON, S. HOWARD BARTLEY, AND ROBERT M. JEWELL | |
| Brightness changes associated with variation in the number of pulses in a repetitive photic train: Sensory implications of the Alternation of Response Theory, II | 107 |
| BY THOMAS M. NELSON, S. HOWARD BARTLEY, AND R. M. JEWELL | |
| Using adverbs as multipliers in semantic differentials | 115 |
| BY ROGER B. ALLISON, JR. | |
| Performance selection through radiation death in rhesus monkeys | 119 |
| BY ROGER T. DAVIS AND JAMES P. STEELE | |
| The correlation of the Porteus Maze and the Gestalt Continuation as personnel selection tests of peripheral peoples | 137 |
| BY A. JAMES GREGOR AND D. ANGUS MCPHERSON | |
| Types of emotions or dimensions of emotion? A comparison of typal analysis with factor analysis | 143 |
| BY ALFRED G. DIETZE | |
| Information "chunking" and short-term retention | 161 |
| BY M. S. MAYZNER AND R. F. GABRIEL | |
| The effect of frequency of presentation on the choice of nonsense syllables | 165 |
| BY JAMES C. BECKNELL, JR., WARNER R. WILSON, AND J. C. BAIRD | |
| The production of random letter sequences in schizophrenics | 171 |
| BY HALIDE S. YAVUZ | |
| Brightness from repetitive photic trains varying in pulse number and separation intervals: Sensory implications of the Alternation of Response Theory, III | 175 |
| BY THOMAS M. NELSON, S. HOWARD BARTLEY, AND ROBERT M. JEWELL | |
| HSPQ norms on a state-wide delinquent population | 185 |
| BY GEORGE R. PIERSON AND ROBERT F. KELLY | |
| Variables affecting report of awareness in verbal conditioning | 193 |
| BY LEONARD KRASNER AND LEONARD P. ULLMANN | |
| The effect of experimentally-induced "awareness" upon performance in free-operant verbal conditioning and on subsequent tests of "awareness" | 203 |
| BY WILL K. WEINSTEIN AND REED LAWSON | |
| The effect of sensory deprivation upon scores on the Wechsler Adult Intelligence Scale | 213 |
| BY MALCOLM H. ROBERTSON AND DOUGLAS J. WOLTER | |
| Size discrimination under two conditions of photic intermittency | 219 |
| BY THOMAS M. NELSON, S. HOWARD BARTLEY, AND ROBERTA F. WISE | |
| Notes concerning "nonmotion" responses while viewing rotating stimulus objects | 227 |
| BY THOMAS MULHOLLAND | |
| Instructional sets and motion perceived while viewing rotating stimulus objects | 233 |
| BY THOMAS MULHOLLAND | |
| A behavioristic rationale for the development of psychosomatic phenomena | 239 |
| BY SHELDON J. LACHMAN | |

THE LEADER BEHAVIOR OF UNITED STATES SENATORS*

Ohio State University and Carnegie Institute of Technology

RALPH M. STOGDILL, OMAR S. GOODE, AND DAVID R. DAY

A. INTRODUCTION

Recent empirical research on leadership has been concerned primarily with studies of small groups and such formally structured organizations as business firms, military organizations, and school systems. Comparatively less empirical research effort has been devoted to another important type of leadership in our society—that of the political leader who is elected by a large number of widely dispersed and loosely organized individuals.

Whereas much of the behavior of the executive is concerned with the structure and function of the organization that maintains and supports his leadership, the political leader must appeal directly to the people whose votes he desires. The latter can depend only to a limited degree upon his party organization. His success in gaining votes is likely to be based upon his persuasiveness and visible representation of the interests of his followers rather than upon behaviors that maintain an effectively functioning formal organization.

The research reported in this paper was designed to determine whether a newly developed set of scales can be used to provide meaningful descriptions of the leader behavior of outstanding political leaders. In accordance with the suppositions advanced in the preceding paragraph, it is hypothesized that high ranking political leaders, such as United States Senators, will be described as high in persuasiveness and representation of followers.

B. THE SAMPLE

Two copies of the Leader Behavior Description Questionnaire were mailed to each senator in the United States Senate, with the request that the questionnaires be given to "two persons who know you well enough to describe you accurately as a political leader in the state that you represent." The describer was asked to identify neither himself nor the senator whom he described. No follow up was used to increase the number of replies.

Forty-four usable questionnaires were returned. This number represents

* Received in the Editorial Office on February 2, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

44 descriptions, not necessarily 44 senators. In regard to political party, 15 questionnaires were checked as descriptions of Republicans, and 29 as descriptions of Democrats. Although it cannot be stated that the sample is strictly representative of the total population of United States Senators, the number of returns is large enough for meaningful analysis.

C. THE QUESTIONNAIRE

For this study, seven new subscales were added to the Leader Behavior Description Questionnaire (1), developed by the Ohio State Leadership Studies Staff. The revision (2) used in this study consisted of 65 items, divided among nine subscales, as follows:

1. *Representation*—speaks and acts as the representative of the follower group (5 items).
2. *Tolerance of Uncertainty*—tolerates postponement and uncertainty of outcome without anxiety (8 items).
3. *Persuasiveness*—presents point of view with conviction, and influences by convincing argument (10 items).
4. *Initiation of Structure*—clarifies own role and lets follower group know what is expected (10 items).
5. *Tolerance of Freedom of Action*—allows follower group scope for initiative in decision and action (5 items).
6. *Role Assumption*—assumes the leadership role; does not surrender leadership to other persons (7 items).
7. *Production Emphasis*—pushes for production and emphasizes results (5 items).
8. *Consideration*—is considerate of the well-being of the follower group (10 items).
9. *Demand Reconciliation*—effectively reconciles conflicting demands (5 items).

Each item in the Leader Behavior Description Questionnaire is answered by circling one of five possible responses: always, often, occasionally, seldom, never. In scoring the questionnaire, each item receives a score value that can range from five to one (or one to five for negatively stated items), as shown in the following examples.

He is a very persuasive talker:

Always (5) Often (4) Occasionally (3) Seldom (2) Never (1)

He fails to take necessary action:

Always (1) Often (2) Occasionally (3) Seldom (4) Never (5)

The score for each subscale consists of the sum of the values received on the items in the subscale.

D. METHOD OF ANALYSIS

The scores of the nine subscales were intercorrelated and factor analyzed, using the method of principal factors. Due to the small number of subscales, it was decided to continue the analysis until all residual entries were reduced to zero. All communalities are necessarily 1.00. The orthogonal factors were rotated by the varimax method. The matrix of intercorrelations, means, standard deviations, and reliability coefficients are shown in Table 1.

TABLE 1
INTERCORRELATIONS AMONG NINE SUBSCALES

| Subscale | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|--------------------------|------|------|------|------|------|------|------|------|------|
| 1. Representation | — | .30 | .63 | .63 | .27 | .64 | .34 | .63 | .44 |
| 2. Tolerates uncertainty | .30 | — | .25 | .21 | .39 | .35 | -.21 | .62 | .64 |
| 3. Persuasiveness | .63 | .25 | — | .30 | .24 | .39 | .25 | .40 | .23 |
| 4. Structure | .63 | .21 | .30 | — | .22 | .59 | .68 | .53 | .44 |
| 5. Tolerates freedom | .27 | .39 | .24 | .22 | — | .16 | -.04 | .50 | .29 |
| 6. Role assumption | .64 | .35 | .39 | .59 | .16 | — | .32 | .62 | .72 |
| 7. Production emphasis | .34 | -.21 | .25 | .68 | -.04 | .32 | — | .20 | .18 |
| 8. Consideration | .63 | .62 | .40 | .53 | .50 | .62 | .20 | — | .77 |
| 9. Demand reconciliation | .44 | .64 | .23 | .44 | .29 | .72 | .18 | .77 | — |
| Mean | 20.7 | 28.2 | 42.5 | 38.8 | 18.3 | 28.8 | 20.6 | 41.1 | 20.7 |
| Standard deviation | 2.5 | 6.1 | 4.6 | 5.5 | 3.1 | 4.0 | 2.6 | 5.9 | 3.5 |
| Reliability | .80 | .83 | .82 | .72 | .64 | .65 | .38 | .85 | .81 |
| Mean (10 items*) | 41.4 | 35.3 | 42.5 | 38.8 | 36.6 | 41.4 | 41.2 | 41.1 | 41.4 |

* Estimated mean for 10 items to correct for unequal numbers of items in the subscales.

It will be noted in Table 1 that the highest mean score (42.5) is obtained on Persuasiveness, as hypothesized. However, the corrected mean for Representation (41.4) is the same as, not higher than, the means for Role Assumption and Demand Reconciliation. The lowest average score (35.3) is on Tolerance of Uncertainty. The means do not differ to a statistically significant degree.

E. RESULTS

The rotated factor loadings are shown in Table 2. Inspection of this table indicates that several subscales have appreciable loadings on Factor I. Role Assumption has a loading of .89 and Demand Reconciliation a loading of .51. Other subscales weighted on this factor are Representation (.30), Initiating Structure (.27), and Consideration (.32). This factor describes a dimension of behavior that is characterized by assumption of leadership, control of conflicting demands, clear definition of role expectations, and representation and consideration of the interests of followers. The combi-

TABLE 2
ROTATED FACTOR LOADINGS

| Subscale | I | II | III | IV | V | VI | VII | VIII | IX | h^2 |
|--------------------------|------|------|------|------|------|------|------|------|------|-------|
| 1. Representation | .30 | .19 | .11 | .38 | .13 | .80 | .17 | .18 | .07 | 1.00 |
| 2. Tolerates uncertainty | .15 | -.14 | .93 | .12 | .20 | .09 | .14 | .05 | .11 | 1.00 |
| 3. Persuasiveness | .13 | .11 | .09 | .95 | .11 | .21 | .07 | .04 | .03 | 1.00 |
| 4. Structure | .27 | .58 | .14 | .08 | .12 | .29 | .13 | .67 | .06 | 1.00 |
| 5. Tolerates freedom | .05 | -.01 | .15 | .09 | .97 | .07 | .11 | .04 | .06 | 1.00 |
| 6. Role assumption | .89 | .18 | .17 | .17 | .03 | .24 | .14 | .15 | .13 | 1.00 |
| 7. Production emphasis | .12 | .97 | -.13 | .12 | -.04 | .09 | .05 | .10 | .03 | 1.00 |
| 8. Consideration | .32 | .13 | .39 | .17 | .31 | .26 | .69 | .13 | .21 | 1.00 |
| 9. Demand reconciliation | .51 | .14 | .47 | .03 | .11 | .11 | .29 | .06 | .63 | 1.00 |
| Fractional contribution | .152 | .156 | .148 | .129 | .126 | .102 | .074 | .060 | .053 | 1.00 |

nation of role assumption, demand reconciliation, and initiation of structure, suggests that the leadership role is under active and firm control. The factor is identified as *Control of the Leadership Position*.

On Factor II, Production Emphasis shows a loading of .97 and Initiating Structure a loading of .58. Here, pushing for production is associated with the clarification of role expectations. The factor appears to merit the title *Facilitation of Productive Outcomes*.

Factor III shows a loading of .93 on Tolerance of Uncertainty. Also loaded on this factor are Consideration (.39) and Demand Reconciliation (.47). Tolerance of uncertainty and postponement is here associated with ability to reconcile conflicting elements and with consideration of the welfare of followers. The factor is identified as *Reconciling Tolerance of Uncertainty*.

Factor IV appears with loadings of .95 on Persuasiveness and .38 on Representation of the follower group. This factor may be regarded as a measure of *Follower Oriented Persuasiveness*.

Factor V, with loadings of .97 on Tolerance of Freedom of Action and .31 on Consideration, is essentially descriptive of *Considerate Tolerance of Follower Freedom of Action*.

Factor VI has a loading of .80 on Representation. Also loaded on this factor are Initiating Structure (.29), Role Assumption (.24), Consideration (.26), and Persuasiveness (.21). This factor appears to merit the title *Active Representation of Followers*.

* In Factor VII, Consideration exhibits a loading of .69 and Demand Reconciliation a loading of .29. Here, consideration of the welfare of followers is associated with reconciliation of conflicting demands. This factor is identified as *Reconciling Consideration*.

Factor VIII is a specific factor, with a loading of .67 on *Initiation of Structure*.

Factor IX, with loadings of .63 on Demand Reconciliation and .21 on Consideration is identified as *Reconciliation of Conflicting Demands*.

F. DISCUSSION

It was found as hypothesized, that United States Senators are described as high in persuasiveness. It was not anticipated that control of the leadership position would emerge as a strongly defined factor, but a bit of reflection suggests that this is a reasonable finding. The elected political leader cannot depend upon a formal organization structure for the maintenance of his position of leadership. He can accomplish this only through an active and continuous assumption of the leadership role.

Two of the factors are defined essentially by a single subscale each. For the remaining seven factors, two or more subscales contribute loadings toward the definition of each factor. The titles of these factors (Control of the Leadership Position, Facilitation of Productive Effort, Reconciling Tolerance of Uncertainty, Follower Oriented Persuasiveness, Considerate Tolerance of Member Freedom, Active Representation of Followers, and Reconciling Consideration) suggest behaviors that should be of value to the political leader who retains his position of leadership, not by daily contact, but by occasional and long-range communications.

G. SUMMARY

A set of newly developed Leader Behavior Description subscales was used to obtain 44 descriptions of the leader behavior of United States Senators. Factor analysis of the subscale intercorrelations produced nine factors. The factors were identified as follows:

- I. Control of the Leadership Position
- II. Facilitation of Productive Effort
- III. Reconciling Tolerance of Uncertainty
- IV. Follower Oriented Persuasiveness
- V. Considerate Tolerance of Member Freedom of Action
- VI. Active Representation of Followers
- VII. Reconciling Consideration

VIII. Initiation of Structure

IX. Reconciliation of Conflicting Demands

The results suggest that the new subscales can be used to obtain meaningful and useful descriptions of the leader behavior of political leaders.

REFERENCES

1. STOGDILL, R. M., & COONS, A. E., *Eds.*, Leader behavior: Its description and measurement. Columbus: The Ohio State University, Bureau of Business Research, Monograph No. 88, 1957.
2. STOGDILL, R. M., GOODE, O. S., & DAY, D. S. New leader behavior description subscales. *J. of Psychol.*, 1962, **54**, 259-269.

Bureau of Business Research

Ohio State University

Columbus 10, Ohio

INTEREST, MEDIA, AND ORDER EFFECTS IN PERSUASIVE COMMUNICATIONS*¹

• *Department of Psychology, Alfred University*

ROBERT E. LANA²

A. PROBLEM

Within the general framework of research done to examine pre-experimental conditions of the subject which influence order effects (primacy-recency) in opinion change, interest of the subject in the topic of the communication has not been systematically studied. Variables such as controversy of the topic (2, 4), familiarity with the topic (1, 4), and awareness by the subject of the intent of the communicator in the opinion change situation (2, 3) have been examined within the context of order effects and opinion change. Presumably, the subject's degree of interest in the topic may represent a variable which is important in determining order effects in opinion change in a manner similar to the degree of controversy of the topic, or his familiarity with it. It is predicted that a significant primacy effect will result under the condition where subjects are greatly interested in the topic of the communication, and a significant recency effect will result under conditions where the subject's interest in the topic is minimal. This prediction is based upon the results of studies by Lana (1, 2) which indicated that subjects highly familiar with the topic, or where they judged the topic to be very controversial, were more influenced by the first argument (primacy), be it pro or con, than by the second argument.

All previous studies in this series (1, 2, 3) and most of those contained in the literature on order effects in persuasive communications utilizing a prose presentation, involved the situation in which someone read the communi-

* Received in the Editorial Office on February 7, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ This study was supported by Research Grant #M-4830 from the National Institute of Mental Health, United States Public Health Service. All materials used in this study are available to research workers only, free of charge, and can be procured by writing to the author.

² The author wishes to thank Ralph L. Rosnow and Gordon Russell for their help in collecting the data.

cations to the assembled group of subjects. It is my contention that the medium of communication may very well play a determining role in producing order effects. Various mass media phenomena have been examined for their effects on public opinion, but little of this has been done within the framework of the laboratory experiment in communication. In this study the usual read communications will be compared to the identical communications (pro and con) presented to the subjects by tape recording for order effects. Hence, in this later situation the communication presentation is depersonalized to some extent, compared with the same communication being read to the subjects by an individual actually present during the experiment.

B. PROCEDURE

One hundred and six students in the beginning psychology courses at The American University were used to form eight unequal groups. All subjects filled out an opinion questionnaire on racial integration of public facilities, such as schools, eating places, swimming pools, etc. This topic had been judged (2) to be of high controversy by subjects of the population from which the current sample was drawn. A questionnaire of the Likert variety (five items, five alternatives), tapping the interest of the subjects in the topic, was also administered at this time. Roughly half of the group, 10 days later, listened to pro and con integration arguments presented by tape recorder. The other half had the arguments read to them by one of the experimenters. Each of these communications is approximately 425 words in length and takes about four minutes to present. These arguments and the questionnaires are described more fully elsewhere (2). The usual pro-con and con-pro orders of presentation of the arguments were also present for both of these groups. The posttest integration questionnaire was administered to all subjects immediately after the presentation of the arguments. In the analysis of the pretest and posttest opinion change scores, the various groups were subdivided into interest groups, high or medium. The groups were formed on the basis of a median split of the interest questionnaire scores given during the pretest. Since very few subjects were at the disinterested end of this questionnaire, the labels "high" and "medium" are appropriate.

C. RESULTS

Media through which the arguments were presented (read or tape), high and medium interest in the topic, and order of presentation of the

arguments were analyzed as main effects. The results of the subtractive procedure follow:

TABLE 1
RESULTS OF SUBTRACTIVE TECHNIQUE

| | A ₁ (tape) | | | | A ₂ (read) | | | |
|-----------|-----------------------|----------------|--------------------|----------------|-----------------------|----------------|--------------------|----------------|
| | B ₁ | | B ₂ | | B ₁ | | B ₂ | |
| | C ₁ | C ₂ | C ₁ | C ₂ | C ₁ | C ₂ | C ₁ | C ₂ |
| Pre | 20.27 | 19.08 | 17.73 | 17.77 | 19.60 | 15.10 | 18.27 | 15.30 |
| Post | 20.40 | 19.38 | 18.93 | 17.15 | 20.07 | 13.30 | 17.40 | 16.40 |
| Gain | +13 | +30 | +1.20 | -.62 | +47 | -1.80 | -.87 | +1.10 |
| | -17 (recency) | | +1.82 (primacy) | | +2.27 (primacy) | | -1.97 (recency) | |
| <i>t</i> | <1 | | 2.64 | | 3.78 | | 2.32 | |
| <i>df</i> | 28 | | 26 | | 28 | | 18 | |
| <i>p</i> | | | <.02 >.01 | | <.01 | | <.05 >.02 | |

Note: B₁ = high interest, B₂ = Medium interest, C₁ = pro-con, C₂ = con-pro.

The analysis of covariance of the data follows:

TABLE 2
ANALYSIS OF COVARIANCE SUMMARY

| Source | <i>df</i> | <i>MS</i> | <i>F</i> | <i>p</i> |
|------------|-----------|-----------|----------|--------------|
| A Media | 1 | 16 | 2.64 | <.20 > .10 |
| B Interest | 1 | 0 | 0 | |
| C Order | 1 | 15 | 2.48 | <.20 > .10 |
| AB MXI | 1 | 0 | 0 | |
| AC MXO | 1 | 95 | 15.68 | <.001 |
| BC IXO | 1 | 8 | 1.32 | > .20 |
| ABC MXIXO | 1 | 63 | 10.40 | <.005 > .001 |
| error | 97 | 6.06 | | |

The media and order interaction effect and the second order interaction are significant. From the subtractive technique, it is evident that the significant interaction effects are reflected in three of four significant *t*-tests, two of which indicate primacy effects and the other a recency effect.

D. DISCUSSION

None of the main effects of the covariance analysis were significant, but the order by media interaction effect and the triple interaction were significant. This latter effect is largely uninterpretable, but leads us to conclude that the factors determining opinion change in this experiment are extremely complex and will eventually have to be re-examined in other designs.

The subtractive technique indicated that significant primacy effects were present for the tape medium-interest group and for the read high-interest group. A significant recency effect occurred in the read medium-interest group. The high-interest tape group showed no significant directional effect. There is no doubt that the media of presentation has an effect on opinion change as a result of this study, but it is not clear exactly how this influence occurs. Order and media have yielded the strongest change effect here. If one examines only the group which was read the communications it is clear that the high-interest group was affected in the direction of primacy, and the medium-interest group in the direction of recency. These results alone are consistent with those on controversy of, and familiarity with the topic (1, 2) if one considers that interest may have the same influential status in changing opinion as familiarity with the topic and degree of controversy. Low familiarity and low controversy groups produced a recency effect, or no effect. A primacy effect was generated from the high familiar group in the above studies. However, when the medium of communication is changed to a more impersonal form (tape) the medium-interest group yields a primacy effect and the high-interest group neither a primacy nor a recency effect. Hence, for medium-interest the effects are reversed. Apparently a tentative conclusion must be that a mode of presentation other than the usual reading of an argument face to face with the experimental group yields significantly different order effects. It is to be noted that all of the studies reported from our laboratory used a read presentation of the communications, except for the present one.

E. CONCLUSIONS

By way of conclusion, it perhaps will be interesting to the reader to examine the results of seven recent studies on primacy-recency order effects with prose material presented as the communication. Table 3 summarizes these seven studies.

There is a similarity in the effects of controversy, familiarity, awareness of manipulation, and interest on order effects in opinion change which seems to indicate that the less exposed group generally yields a recency effect, while high exposed groups yield primacy effects. Admittedly, there is difficulty in attempting to explain the "no effect" results. However, with data of this sort, where precise quantification is not possible nor necessarily desirable, the systematic repetition of certain phases of various experiments can be illuminating. The fact that almost twice as many primacy effects were

TABLE 3
SUMMARY OF RECENT PRIMACY-RECENCY STUDIES

| Primacy | Recency | No Effect |
|---|-----------------------|--|
| 1. (Lana, 1962) High familiarity | Low familiarity | Medium familiarity |
| 2. (Lana, 1962) High controversy | | Low controversy |
| 3. (Lana, 1962) Medium controversy, hidden pretest | | High controversy, hidden and unhidden pretests Medium controversy, unhidden pretest |
| 4. (Lana, 1962) Tape, medium interest Read, high interest | Read, medium interest | Tape, high interest |
| 5. (Schultz, 1962) Moderately aware (of intent of communication) | Unaware | Strong awareness |
| 6. | | (Thomas, Webb, & Tweedie, 1961) High controversy with high and low familiarity |
| 7. | | (Lana, 1962) (replication of 2 above) High controversy Low controversy |

isolated as recency effects is significant because seven distinct experiments, performed under the direction of three different investigators, are involved. The theoretical interpretation of these results awaits further study, however.

REFERENCES

1. LANA, R. E. Familiarity and the order of presentation in persuasion. *J. Abn. & Soc. Psychol.*, 1961, **62**, 573-577.
2. ———. Order effects in persuasive communications. Progress Report #1, Research Grant #M-4830, National Institute of Mental Health, United States Public Health Service, May, 1962.
3. SCHULTZ, D. P. Primacy-recency as a function of time and of levels of awareness of manipulatory intent. Unpublished Doctoral dissertation, The American University, 1962.
4. THOMAS, E. J., WEBB, S., & TWEEDIE, J. Effects of familiarity with a controversial issue on acceptance of successive persuasive communications. *J. Abn. & Soc. Psychol.*, 1961, **63**, 656-659.

Department of Psychology
Alfred University
Alfred, New York

EFFECTS OF MASCULINE-FEMININE EGO-INVOLVEMENTS ON THE ACQUISITION OF A MIRROR-TRACING SKILL*

Gustavus Adolphus College

DAVID NOER AND JAMES O. WHITTAKER

A. PURPOSE

Previous research has shown that verbal learning and memorization can be significantly improved when a subject is ego-involved (1). There has, however, been little work done on the effects of ego-involvement in the acquisition of motor skills, or on the effects of different types of ego-involvement in the acquisition of such skills.

In this experiment, the problem involves assessing the effects of ego-involvement in the acquisition of a motor skill, and also the effects of feminine ego-threats vs. masculine ego-threats in the acquisition of such skills. Our first hypothesis is that ego-involvement will lead to significant improvement in the acquisition of a mirror-tracing skill. The second hypothesis is that masculine ego-threats will result in greater ego-involvement than feminine ego-threats, and that such differences will be reflected in higher task scores for the males.

B. METHOD

Twenty-eight undergraduate students were divided into two groups of seven males and seven females each. Group A, the control group, did thirty mirror star tracings after being told to do their best at the task. A short rest period intervened between the first and last fifteen trials. Group B, the experimental group, performed the first fifteen tracings under the same conditions as subjects in Group A. Following this, an attempt was made to ego-involve subjects in the experimental group. Male subjects were individually told that:

You have now finished fifteen trials and are half finished with the experiment. You can take a short break before you begin the last half. I suppose you are wondering about the significance of this experiment. I don't see why I couldn't explain the purpose of it to you. Recently a report was published in a psychological journal, stating that when using

* Received in the Editorial Office on February 8, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

this very same test, boys did nearly twice as well as girls. The article offered no reason for this, but definitely concluded that males did much better than females. Actually what we are doing here, in this experiment, is testing the validity of this conclusion.

We have found in running this test that the conclusions published are true. Most boys here do at least twice as well as girls.

At this point the experimenter would thumb through the first fifteen tracings done by the subject and then make the following comments:

On the basis of your first 15 tracings, you seem to be working about on the norm for girls, falling below in some cases. Most of the other males in our sample have done about twice as well as you have done. I am sure that you can do better, so please do the best you can.

Female subjects were told the reverse of this, i.e., that female subjects were superior to males, but that in this case, the girl's score was at or below that of male subjects.

The apparatus used was a modification of the mirror tracing board described by Townsend (2). The base of the tracing board was 13" \times 10" and the mirror was seven inches high and five inches wide. A six pointed star was used for subjects to trace. The distance between opposite points was five and one-quarter inches, from outside edge to outside edge. The inner edge of the star was three-sixteenths of an inch inside the outer edge, and the area between the lines was divided into 72 segments or boxes to facilitate scoring.

C. RESULTS

Each star traced by a subject was scored as follows: each of the 72 segments or boxes in the star was checked for errors. If the pencil left the paper in the box or if it touched one of the edges outlining the star, the box was counted as an error. The total number of boxes in which errors occurred was subtracted from 72. The remaining number was then multiplied by 60, to give the score, or rate per minute done without error. A mean score (mean rate per minute) was calculated for the first 15 and last 15 trials separately. The overall means for both control and experimental groups, differentiated according to sex, are shown in Table 1.

Tests of the differences between group means by sex, on trials 1-15 were not significant. The difference between means for males in the two groups on trials 16-30, on the other hand, proved significant at the .01 level of confidence. Similarly, a test of the difference between means for females on trials 16-30 in the two groups, proved significant ($p = < .02$). A test

TABLE 1
MEAN SCORES BY SEX FOR BOTH GROUPS

| Trials | Control group | | Experimental group | |
|--------|---------------|--------|--------------------|--------|
| | Male | Female | Male | Female |
| 1-15 | 72.71 | 96.03 | 69.32 | 87.25 |
| 16-30 | 119.12 | 143.08 | 175.99 | 200.15 |

of the difference in improvement between trials 1-15 and trials 16-30 for subjects in the experimental group, by sex, proved to be not significant.

D. DISCUSSION

The results of the experiment substantiate the hypothesis that ego-involvement in general, results in significant improvement in the acquisition of a mirror tracing skill. Both males and females, when ego-involved, perform at significantly higher levels than do subjects who are not ego-involved.

The second hypothesis, stating that masculine ego-threats yield greater ego-involvement than feminine ego-threats, was not verified. There are several possible explanations of these results. First, the hypothesis may be inaccurate. Second, it is possible that the hypothesis is accurate, but our procedure was inappropriate for testing it. Third, the procedure may have "over ego-involved" the male subjects resulting in a performance decrement. Fourth, the hypothesis may be inaccurate for the type of subjects used in this experiment. Thus masculine ego-threats might result in greater ego-involvement among lower socioeconomic group males, relative to females, than among those males from upper-middle and upper-class socioeconomic groups. There is some evidence which suggests that this latter hypothesis is correct. Masculine and feminine social roles appear to be more distinct or more clearly differentiated among those in lower socioeconomic groups. Secondly, concern with masculinity appears to be of greater consequence to lower socioeconomic males than to upper-middle and upper-class males. If this be true, an experiment such as the one reported here, conducted with subjects from lower socioeconomic groups, should yield results reflecting this difference.

REFERENCES

1. SHERIF, M., & CANTRIL, H. *The Psychology of Ego-Involvements*. New York: Wiley & Sons, 1947.
2. TOWNSEND, J. *Introduction to Experimental Method*. New York: McGraw-Hill, 1953.

Department of Psychology
Gustavus Adolphus College
St. Peter, Minnesota

A FURTHER STUDY OF DETERMINANTS OF PHENOMENAL DISTANCE IN PLANE TARGETS PERCEIVED AS THREE-DIMENSIONAL SCENES*

Department of Psychology, Michigan State University

JANE E. RANNEY AND S. HOWARD BARTLEY

A. PROBLEM

Phenomenal distance is the apparent nearness of an object. Some objects, though located at the same metric distance, are seen as nearer or farther than others. Previous studies have investigated the effects of viewing distance (3, 7, 9, 10, 11), print size (2, 3, 12), and asymmetry of scenes (1, 4, 6) on apparent distance. However, the effects of item position, crucial item size, and location of a background item have not been clearly determined. Four hypotheses concerning the effect of these factors on phenomenal distance of items in photographed scenes and their interrelations were investigated in this experiment.

First we hypothesized that large items seen at a given distance appear nearer than small items located at the same position in the scene. Second, the left-right imbalance is dependent upon the background item appearing on the right, or at least the maximum effect occurs only when the background item is on the right. Third, the relative position of the large background items has a greater effect on the phenomenal distance of small items than of large items. Fourth, the difference between the phenomenal distance of items on the right and left is greater for large than for small items.

B. METHOD

1. *Observers*

Twelve male students in an introductory experimental psychology course at Michigan State University served as observers. They all met a visual acuity criterion of 20/20 or better in the right eye.

2. *Apparatus*

Five photographs of an artificial scene were printed in black and white on glossy paper, and the negatives simply inverted to produce mirror images.

* Received in the Editorial Office on February 18, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

These provided all possible combinations of two item sizes, right and left position, and two conditions of background complexity.

Trees, a hill and road provided a complex background on the right of the scene, and the left or simple portion of the background was relatively flat with a lake on the horizon. The crucial item, either a small square black block or one of the same width but five times taller, was to the left or right of center in the midground for the variable targets. The small block was in the center for the standard targets. Shadows in the photographs were minimized by taking the photographs in diffuse overhead lighting. Sample photographs (Figure 1) show how realistic the scene was.

In describing the targets, R and L refer to the position of the trees in the background. The tall block is coded as *r* and *l*, and the small black as (*r*), (*c*), and (*l*). Thus in R*l* the trees are on the right, and the tall block on the left.

Eight large (8×10 -inch) prints were used as variable targets. Four of these (R*r* and R*l* and their mirror images, L*l* and L*r*) contained the tall block, and four, R(*r*), R(*l*), L(*l*) and L(*r*), the small block. These variable targets were compared with the two smaller (4×5 -inch) targets, R(*c*) and L(*c*), held at a fixed distance. Since the small block was in the center in both standard targets, in describing these targets only the position of the trees in the standard target will be stated, followed by the comparison target: thus, L*R*(*l*).

The apparatus for presenting targets consisted of an adjustable carriage mounted on a calibrated track 275 inches long; a stationary target holder to the right of this, and a chin rest and blind to block vision in the left eye. By turning a crank on the left of the apparatus, *O* moved the large print along the track. The track was illuminated by diffuse overhead lighting. Targets were seen against a flat black background.

3. Procedure

*O*s were first tested for visual acuity using a "tumbling E" acuity chart. Only individuals with 20/20 acuity or better in the right eye served as *O*s.

To eliminate *O*s who did not necessarily perform the task, a further criterion was established. In a practice series each of six pairs of targets similar but not identical to the experimental targets was matched four times in alternating ascending and descending order.

One of the small practice targets was placed on the right, 29 inches from *O*'s right eye, and a large (variable) target randomly positioned along the calibrated track. *O* was seated before the apparatus and instructed:

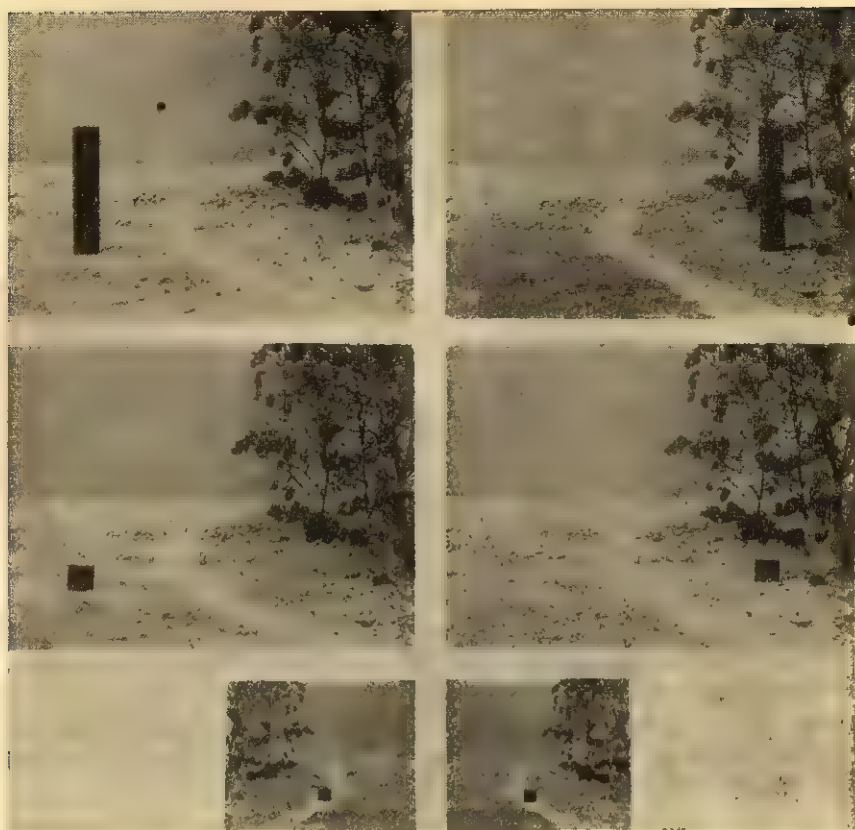


FIGURE 1

SAMPLE TARGETS

The upper targets are Rl and Rr for the left and right respectively. Those in the center contain the small block and are coded R(l) and R(r). The lower targets, coded L(c) and R(c), served as standard targets. Mirror images of the upper four targets were also used.

This is an experiment in distance judgment. Rest your chin on the chin rest. Now you should be able to see the pictures only with your right eye. I want you to move the large picture with this crank until this object (the block) looks as far away from you as the one in the small picture. The objects should look the same *distance* from you, not the same size. I'll give you several practice trials, so go ahead and try it.

Os were allowed to repeat a trial if they were quite dissatisfied with the response, and were free to take a break whenever they wished.

The variable target was placed at some random point along the track and *O* moved the target toward or away from himself until the block appeared to be at the same distance as a similar block in the standard target. The measure then taken was the distance in inches from *O* to the variable target. A target with an item which appears relatively near is placed farther from *O* than a target containing an item which appears farther away.

Unless *O*'s range on three of these six practice conditions was 10 inches or less, he did not complete the experiment. This criterion eliminated one *O*. From this practice *O*s developed a more stable criterion of equality of distance, and therefore a more stable range for the experimental series.

In the main experiment, the same instructions and procedure were used as in the practice series. The sixteen pairs of targets were presented in random order to each of twelve *O*s. Since *O*s made two ascending and two descending matchings for each pair, they made a total of 64 comparisons in addition to the 24 practice trials.

C. RESULTS AND DISCUSSION

The four readings taken under each condition were averaged to provide scores for the observers. Results of an analysis of variance on these scores are presented in Table 1. Results of a Bartlett's test for homogeneity of variance of the triple interactions involving observers, and of the four- and five-way interactions involving observers did not permit rejection of the hypothesis of homogeneity of variance. For the triple interactions χ^2 was 4.2906, $df = 5$, $p > .75$. For the four- and five-way interactions, $\chi^2 = 4.3894$, $df = 4$, $p > .75$. Therefore, pooled variances were computed to be used as error terms for main effects and double interactions involving observers (pooled three-way interaction variance), and triple interactions involving observers (pooled four- and five-way interaction variance). In all other cases the error term was the interaction of the source of variance being tested with observers. For example, the error term for Block Size was the Block Size by Observers interaction mean square.

The results tend to confirm three of the four hypotheses proposed.

Hypothesis 1. It was expected that large items at a given distance would appear nearer, and therefore these prints would be placed farther along the track than prints depicting small items located at the same position in the scene. This is true for seven of the eight comparisons, as is apparent from Figure 2A. The mean distance for prints containing large blocks was 64.78, for small, 62.68. The analysis of variance showed the effect of Block Size to be significant at nearly the .01 level of significance.

TABLE 1
SUMMARY OF ANALYSIS OF VARIANCE BASED ON SCORES FOR EACH O
UNDER EACH CONDITION

| Source | df | | |
|--|-----|----------|-----------|
| A Tree position (standard) | 1 | 35.6644 | 3.1443 |
| B Tree position (variable) | 1 | 17.9769 | 1.9374 |
| C Block position | 1 | 23.2060 | 1.5741 |
| D Block size | 1 | 210.9456 | 9.2491* |
| E Observers | 11 | 803.0905 | 85.1859** |
| A × B | 1 | 1.2113 | .1572 |
| A × C | 1 | 39.6488 | 5.6939* |
| A × D | 1 | 15.6124 | 1.9805 |
| B × C | 1 | 217.2816 | 12.1568* |
| B × D | 1 | 52.3441 | 5.4725* |
| C × D | 1 | 2.6958 | .4100 |
| A × E | 11 | 11.3426 | 1.2931 |
| B × E | 11 | 9.2787 | .9842 |
| C × E | 11 | 14.7423 | 1.5637 |
| D × E | 11 | 22.8071 | 2.4192* |
| A × B × C | 1 | 2.8154 | .2659 |
| A × B × D | 1 | .9143 | .0926 |
| A × C × D | 1 | .0731 | .0092 |
| B × C × D | 1 | 38.7450 | 5.5985* |
| A × B × E | 11 | 7.7063 | 1.0017 |
| A × C × E | 11 | 6.9634 | .9051 |
| A × D × E | 11 | 7.8829 | 1.0246 |
| B × C × E | 11 | 17.8732 | 2.3231* |
| B × D × E | 11 | 9.5649 | 1.2432 |
| C × D × E | 11 | 6.5742 | .8545 |
| A × B × C × D | 1 | 4.6098 | 1.4614 |
| A × B × C × E | 11 | 10.5889 | |
| A × B × D × E | 11 | 9.8739 | |
| A × C × D × E | 11 | 7.9291 | |
| B × C × D × E | 11 | 6.9206 | |
| A × B × C × D × E | 11 | 3.1544 | |
| Total | 191 | | |
| Pooled variance, triple interaction involving E | 66 | 9.4275 | |
| Pooled variance, four- and five-way interactions involving E | 55 | 7.6934 | |

* $p < .05$.

** $p < .01$.

Bartley's (2) research showed that the absolute size of an item determines the phenomenal distance. The present research shows that an item which is large relative to the background is also seen as nearer. This confirms Bartley and DeHardt's (5) results.

Hypothesis II. We predicted that the left-right imbalance is dependent on

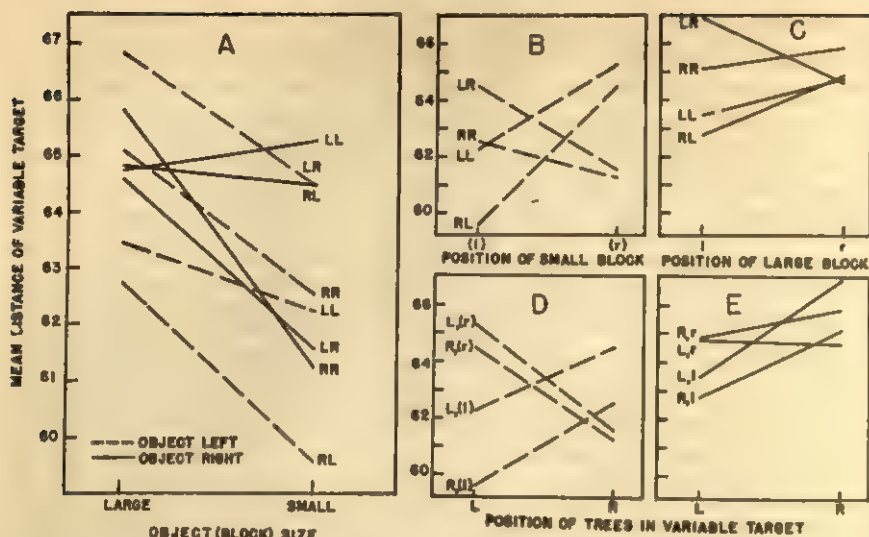


FIGURE 2

Graph A shows the relationship between block size and distance of the variable target. B graphs distance as a function of lateral position of the small block, and C, of the large block in the variable target. D and E present distance as a function of lateral position of trees in the variable targets with the small and large blocks, respectively. Note that B and D refer to targets with the small block, and C and E, with the large block. Mean distance of the variable target is related to phenomenal distance in such a way that the target is placed near *O* when the block looks far away.

the background item appearing on the right. That this is true for the small block is apparent from Figure 2B. However, with the large block this does not appear to be the case. (See Figure 2C.) The item on the left appears nearer than that on the right only in comparing LRl and LRr. This difference in the effect of the position of the background item on the phenomenal distance of the different sized objects is substantiated by the significant Tree Position (Variable) by Block Position by Block Size interaction ($p < .05$), and by the significant Tree Position (Variable) by Block Size ($p < .05$), and Tree Position (Variable) by Block Position ($p < .01$) interactions.

Using photographs from the same series as in this study, Bartley and DeHardt (5) found that with the trees on the right the smaller block appeared nearer on the left than on the right. Our data agree with this. However, with the trees on the left, this study found that the item on the right looked nearer, whereas Bartley and DeHardt found it looked farther, although their differences were not statistically significant.

Hypothesis III. We predicted that the position of the large background item (trees) would have a greater influence on the phenomenal distance of the small block than on the large block. That is, we predicted that the difference between the phenomenal distance of the small block on the left and right depends on the position of the background item rather than of the small crucial item. Evidence for this hypothesis is provided by a comparison of Figures 2B and 2D. In Figure 2B the distance is plotted as a function of the right or left position of the small block in the variable target. In Figure 2D, distance is plotted as a function of the position of the trees in the variable target. From the similarities of these two graphs it is apparent that the phenomenal distance of the small item is largely determined by the position of the large background item.

One would expect, on the other hand, that the phenomenal distance of the larger blocks would not be as greatly determined by the large background items. This is apparent from the differences between Figures 2C and 2E, which are comparable to 2B and 2D, but for the large blocks. Here again, the significant Tree Position (Variable) by Block Size, and Tree Position (Variable) by Block Position by Block Size interactions substantiate the graphical data.

Hypothesis IV. We predicted that left-right differences in phenomenal distance would be greater for large items than for small. There is no direct evidence for this hypothesis since we found that the position of the background item largely determines the phenomenal distance of the small item. Therefore the large and small crucial items cannot be directly compared. The Block Position by Block Size interaction predicted by the hypothesis was not significant. This hypothesis should be investigated in an experiment where the problem is not complicated by the presence of a large background item.

The results of this investigation, in conjunction with those of Bartley and DeHardt (5), have demonstrated the importance of background items in determining phenomenal distance. In contrast, Teichner, Kobrick and Dusek (10) concluded that terrain had little or no effect on distance matches in field situations. Teichner, Kobrick and Wehrkamp (11) found that the linear threshold of equality varied slightly with the terrain. Although the differences between terrains were significant, they were small and not consistent. However their background differences were rather minor, simply the differences in texture of the ground, as silt, sand, or macadam road. So perhaps, as may be inferred from Gogel's (8) work with illusions, phenomenal distance depends on the size of the irrelevant item relative to the size of the crucial item.

Not only does the presence of a large background item affect the phenomenal distance of a crucial item, but the distance is also dependent on the position of the background item as well as its position relative to the size and position of the crucial item.

In conclusion, the results of this experiment show that phenomenal distance of an item in a scene is affected by (a) the size of the item; (b) the lateral position of the item; (c) the lateral position of a relatively large item in the background; and (d) that these factors interact to determine phenomenal distance.

D. SUMMARY

This study investigated the effects of item position, item size, and location of a large background item, on the phenomenal distance of an item in a photographed scene.

Twelve observers compared eight large variable targets to two smaller, fixed targets, matching them so that the crucial items in the photographs appeared to be the same distance from the observer. The variable targets were photographs of an asymmetrical artificial scene containing either a small or large item to the left or right of center, and mirror images of these prints. In the standard targets the small item was in the center of the scene.

The results led to the following conclusions: (1) Large items appear nearer than small items located at the same position in the photographs. (2) The position of a large background item is important in determining the phenomenal distance of the crucial item. It is more important where the crucial item is small, less important when it is large. (3) The item on the left appears nearer than the one on the right when the large background item is on the right.

REFERENCES

1. ADAIR, H., & BARTLEY, S. H. Nearness as a function of lateral orientation in pictures. *Percept. & Motor Skills*, 1958, 8, 135-141.
2. BARTLEY, S. H. Some comparisons between print size, object position and object size in producing phenomenal distance. *J. of Psychol.*, 1959, 48, 347-351.
3. BARTLEY, S. H., & ADAIR, H. J. Comparisons of phenomenal distance in photographs of various sizes. *J. of Psychol.*, 1959, 47, 289-295.
4. BARTLEY, S. H., & DEHARDT, D. C. A further factor in determining nearness as a function of lateral orientation in pictures. *J. of Psychol.*, 1960, 50, 53-57.
5. ———. Phenomenal distance in scenes with independent manipulation of major and minor items. *J. of Psychol.*, 1960, 50, 315-322.
6. BARTLEY, S. H., & THOMPSON, R. A further study of horizontal asymmetry in the perception of pictures. *Percept. & Motor Skills*, 1959, 9, 135-138.

7. DUSEK, E. R., TEICHNER, W. H., & KOBRICK, J. L. The effects of the angular relationships between the observer and the base-surround on relative depth-discrimination. *Amer. J. Psychol.*, 1955, **68**, 483-543.
8. GOGEL, W. C. Perception of the relative distance position of objects as a function of other objects in the field of view. *J. Exper. Psychol.*, 1954, **47**, 335-342.
9. SMITH, O. W. Comparison of apparent depth in a photograph viewed from two distances. *Percept. & Motor Skills*, 1958, **8**, 79-81.
10. TEICHNER, W. H., KOBRICK, J. L., & DUSEK, E. R. Commonplace viewing and depth discrimination. *J. Opt. Soc. Amer.*, 1955, **45**, 913-920.
11. TEICHNER, W. H., KOBRICK, J. L., & WEHRKAMP, R. F. The effects of terrain and observations distance on relative depth discrimination. *Amer. J. Psychol.*, 1955, **68**, 193-208.
12. THOMPSON, R. W., & BARTLEY, S. H. Apparent distance of material in pictures associated with higher order meanings. *J. of Psychol.*, 1959, **48**, 353-358.

Department of Psychology
Michigan State University
East Lansing, Michigan

THE EFFECT OF STIMULUS COLOR AND CONTENT UPON SCHIZOPHRENICS*¹

Department of Psychology, Mills College

ALINE H. KIDD²

A. INTRODUCTION

Modern research demonstrates that perception does not depend solely upon the stimuli which impinge upon the sense organs. The Ames perceptual demonstrations (2) show that differing stimuli may be perceived identically and Hastings (13) concludes that specific external stimuli do not necessarily produce specific perceptions. Research into the determinants of perception has made clear that social and individual past experiences along with their resultant individual value system and idiosyncratic personality organization contribute to the stimuli in an individual's achievement of a percept. Dennis (8) and Piaget (23) stressed the fact that perception must be studied in terms of settings which are socially significant to the perceiver. Helson (15), Hastings (13), Ittelson and Cantril (18) who emphasize the importance of the individual, held that, although experience is ordered, perception is not ordered in advance but is organized creatively by the individual, differing from person to person with the same stimuli and also changing for an individual from occasion to occasion. Blake, Ramsey, and Moran (4) maintain that the individual's personal organization of meaningful experiences affects perceptual activity. The results of the majority of this research indicate that, although there is a general correspondence between the external world and the perceptions of the normal subject, this correspondence is not complete. Bronfenbrenner (5) concluded that personality is a hierarchy of psychological situations which are dispositions to respond in a particular way to selected aspects of the field. Imboden (17) shows that anxiety produces perceptions which are non-veridical or unrealistic in terms of the stimulations. Beach's (3) review of the literature shows that physiological perception studies indicate that

* Received in the Editorial Office on February 28, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ This research was supported by grant #60-3-2 (Q-5) from the California State Department of Mental Hygiene.

² The author is indebted to Russell Gould, M.D. and John Marchesi, M.D., Agnews State Hospital, who provided the schizophrenic subjects and to all department heads at Agnews State Hospital who provided the normal subjects.

changes in body chemistry produce radical alteration in perception. Semi-starvation, extreme vitamin deficiency, metabolic dysfunction, anoxia, drugs and hormonal shifts may all affect perception. In addition, the drugs which produce psychopathological states appear to change perception (1).

There has been some specific research on the relationship between personality and perception. Hastings (13) showed that individuals who were insecure, as measured by scores on the Knutson personality inventory, tended to bring moveable visual stimuli significantly closer to them than did more secure subjects. Smith (27, 28) using a twelve-foot table with parallel viewing lanes and a moveable projected image, found that more secure individuals, selected by their scores on the Bell adjustment inventory and the Knutson personality inventory, tended to bring human faces closer to them than did insecure subjects. He hypothesized that insecure individuals had suffered from a greater number of unpleasant experiences with other people and, therefore, wanted to place themselves farther from other people than did secure subjects.

There have been some studies of schizophrenic perception. Wertheimer (1957) showed that schizophrenics showed smaller kinaesthetic figural after-effects than did normal subjects. Philip (21) showed that certain stimulus figures were reversed less frequently by schizophrenic than by normal subjects. Rausch (24) showed paranoid schizophrenics demonstrated more size constancy than did normals. Lovinger (19) showed that when maximal distance cues or no distance cues were given there was no difference between good-contact schizophrenics, poor-contact schizophrenics, and normals. However, when minimal cues to distance were given, the good contact and control groups did not differ from one another, but poor-contact schizophrenics showed less constancy. Winder (31) suggested that schizophrenics who are still adapting must do so by excessive control.

There have also been some studies of the way in which personal significance may cause deviations in responses to some stimuli. Hindley (16) has shown that schizophrenics are more influenced in their perception by value variables than are normal subjects and less influenced by frequency variables. Dunn (9) showed pictures of scolding, whipping, feeding and neutral scenes and asked his subjects to determine whether the one picture was identical to the other. Some of the pictures were identical and others were progressively more dissimilar. Schizophrenics had difficulty discriminating pictures of the scolding scene, tended to have difficulty with the whipping scenes, but were not different from the normal group on food and neutral scenes. Davis and Harrington (7) found that schizophrenics, matched

with a normal group for performance on stimulus cards containing non-human content, were inferior when tested on cards containing human content.

There have been, however, no systematic studies of schizophrenic responses to a variety of stimuli differing in color and in subject matter. If in such an investigation schizophrenics were found to differ from normals in their responses to a wide range of stimuli, a total perceptual-cognitive defect would be indicated. Goldiamond (12) has indicated that schizophrenics show both normal and abnormal discriminative behaviors. Therefore, it is hypothesized that schizophrenics will *not* display such a deficit. If some stimuli were found to produce responses in schizophrenics which did not differ from those of the normal group and others to lead to abnormal reactions, it would tend to support either a theory like that of Wechowicz (30) of a breakdown of a biochemical reaction under stress, or a theory of a learning situation in which responses to some stimuli have been punished or nonrewarded, and responses to others have been rewarded or not punished, or that some types of responses have been reinforced and others partially or totally extinguished.

The present research was designed to determine whether the responses of schizophrenic patients differ from those of normal subjects in general or only in relation to specific stimuli. If they differ only when the schizophrenic is confronted with specific stimuli, as is hypothesized, the present research will indicate some of the types of stimuli which elicit abnormal responses.

B. METHOD

1. Subjects

The experimental group was composed of 50 schizophrenic patients (25 male and 25 female) who were tested within 48 hours of their admission to a state hospital. They had been drug and shock-free for at least six months and, believing that their performance on the experiment was part of the psychiatric examination to determine whether or not they would be institutionalized, they were highly motivated.

The control group consisted of 50 hospital employees (25 male and 25 female) individually matched with the patients for age within one year, and amount of education within one year. Only employees whose history gave no indication of psychopathology or antisocial behavior were included.

All subjects had normal visual acuity and color vision as indicated by the Snellen charts and by the pseudo-isochromatic plates.

2. Apparatus

A modified thereness-thatness table 16 feet long was set up with two parallel viewing lanes separated by a plywood panel 18 inches high, so that

monocular viewing of each lane was assured. Into the left lane 7-1/4 inches from the outer edge of the table were set five lucite rods at right angles to the table surface and lighted from below by 7-1/2 watt bulbs which were shielded by a panel to avoid extraneous light leakage into the room. The first rod was 45 inches from the subject's eye, and the other rods were spaced at 24 inch intervals beyond the first. In the right lane were tracks on which a grooved-wheeled cart was mounted so that it could be moved easily back and forth with a rope-and-pulley arrangement. The cart contained an enclosed projector and screen unit 31-3/4 inches long, 16-1/4 inches wide and 24 inches high. It was so constructed that only the projected image was seen by the patient on the front-facing surface of the screen. The subject was able to move the unit with the projected image closer to him or farther away from him by rotating the crank-and-pulley either clockwise or counter clockwise. Some slippage built into the crank-and-pulley arrangement by means of medium-duty springs, prevented the subject's aligning the image with the light rods by merely counting the number of turns of the crank.

The slides to be projected contained four types of subject matter: (a) human; (b) animal; (c) food; and (d) nature. Each image appeared in natural, pleasant color; in unnatural, subjectively unpleasant color as judged by subjects in the pilot study; and in black and white. There were two illustrations of each type of subject matter, making at total of 24 images. The range of difference in brightness of the slides was from 13.0 to 13.4 foot candles, as read by a cadmium sulfate meter above a fluorescent light with a ground glass diffusion slide under the image and an opal glass over the image. The results of a pilot study indicate that this amount of difference in brightness does not influence the results significantly on this task.

3. Procedure

Subjects were comfortably seated in front of the equipment with their chins on the padded chinrest. The instructions given were:

In the box at the end of the table there is a projector containing slides. I will show you these slides and ask you to line up the projector box, which moves by turning this crank (indicating crank), up with the second light on the other side. The first light is the one nearest to you. This is light Two, Three, Four and Five (pointing to each rod in turn, starting with the light closest to the subject). After you have lined up the box with the light, I will check on the tape along this side and write down where you placed the box. During the experiment, I want you to keep your chin on the chinrest (indicating) so that you see the lights with your left eye, and the slides with your right eye. Now, try using

the crank. This experiment has to be run in the dark. After I turn out the lights, we will wait for a few minutes until your eyes get used to the dark. When your eyes are used to the dark, I will plug in the equipment and you will be able to see the light and the projector. Have you any questions?

The lights were then extinguished and 10 minutes were allowed for dark adaptation. During this 10 minutes, the subject and experimenter discussed topics which were not related to the experiment. The projector unit was then moved to the end of the table farthest from the subject and the equipment was plugged in. The slides were presented in random order without replacement and the subjects were asked to line the projected images up with the 2nd light. After the subject had lined up each image, the projector unit was moved to the far end of the table and the slide was changed. When all 24 images had been aligned with the required light, starting from the far end of the table (descending series), the room lights were turned on and the subject was given a 10-minute rest period. After another 10-minute period of dark adaptation, the procedure was repeated except that the projector unit was placed at the table's end closest to the subject before each attempted alignment (ascending series).

C. RESULTS

Because the analysis of variance showed no significant difference between the groups, it is clear that schizophrenics and normals respond similarly to some projected images and that, therefore, a total cognitive-perceptual defect was not demonstrated.

For both the Descending Series (screen adjusted from as great a distance from the subject as possible) and the Ascending Series (screen moved from the end of the table closest to the subject), there was a significant difference in responses to types of subject matter (human, animal, nature and food). The application of the t test showed that food images were placed significantly closer to the subject ($t = 2.30$, $P < .02$) than nature images, than human images ($t = 2.08$, $P < .05$) and than animal images ($t = 3.21$, $P < .01$). Animal images were placed significantly farther away from the subject than food images ($t = 2.98$, $P < .01$), nature images ($t = 1.99$, $P < .05$), and human images ($t = 2.51$, $P < .02$).

For both Series, there was a significant difference among the responses to the types of color of the images presented. Although there was no response difference indication between the pleasant and unpleasant color images, the black and white images were placed significantly closer than the pleasant

TABLE 1
ANALYSIS OF VARIANCE OF ASCENDING AND DESCENDING SERIES

| | Source | Sum Square | df | Mean Square | F |
|--------------------------|--------|------------|----|-------------|--------|
| <i>Descending Series</i> | | | | | |
| Groups | A | 946.92 | 1 | 946.92 | 1.88 |
| Subject matter | B | 5,869.68 | 3 | 1,956.56 | 3.89* |
| Color | C | 8,636.46 | 2 | 2,318.23 | 4.62* |
| Interactions | ab | 5,341.42 | 3 | 1,780.47 | 3.54* |
| | ac | 4,798.34 | 2 | 2,399.17 | 4.77** |
| | bc | 564.58 | 6 | 94.10 | .19 |
| | abc | 3,349.12 | 6 | 558.19 | 1.11 |
| | error | 38,226.48 | 76 | 502.98 | |
| | total | 67,733.00 | 99 | | |
| <i>Ascending Series</i> | | | | | |
| Groups | A | 469.92 | 1 | 469.92 | .96 |
| Subject matter | B | 5,888.70 | 3 | 1,962.90 | 4.01* |
| Color | C | 4,620.88 | 2 | 2,310.44 | 4.72* |
| Interactions | ab | 4,303.20 | 3 | 1,434.40 | 2.93* |
| | ac | 3,299.12 | 2 | 1,649.56 | 3.37* |
| | bc | 117.36 | 6 | 19.56 | .04 |
| | abc | 4,754.82 | 6 | 792.47 | 1.62 |
| | error | 37,202.00 | 76 | 489.50 | |
| | total | 60,656.00 | 99 | | |

* .05.

** .01.

($t = 2.28$, $P < .05$) and the unpleasant ($t = 2.16$, $P < .05$) color images.

The ab interaction was significant ($P < .05$) for both Series. Although there was no significant difference between schizophrenic and normal placement of food or animal images, schizophrenics did place human images significantly farther from them ($t = 3.21$, $P < .01$) and nature images significantly closer to them ($t = 2.59$, $P < .01$) than did normal subjects. These data are summarized in Figure 1.

The ac interaction showed similar significance. There was no significant difference in the placement of the pleasant color, the unpleasant color and the black and white images by the normal subjects. However, the schizophrenics placed the black and white images significantly closer to them than the pleasant colors. ($t = 2.41$, $P < .02$) and the unpleasant color ($t = 3.21$, $P < .01$) images, and placed the pleasant color images ($t = 2.05$, $P < .05$) closer to them than the unpleasant color images. These data are summarized in Figure 2.

The correlation between the Descending Series and the Ascending Series for the normal subjects was .78 with a standard deviation of .01; for the schizophrenic subjects was .65 with a standard deviation of .01, and for all subjects was .67 with a standard deviation of .01.

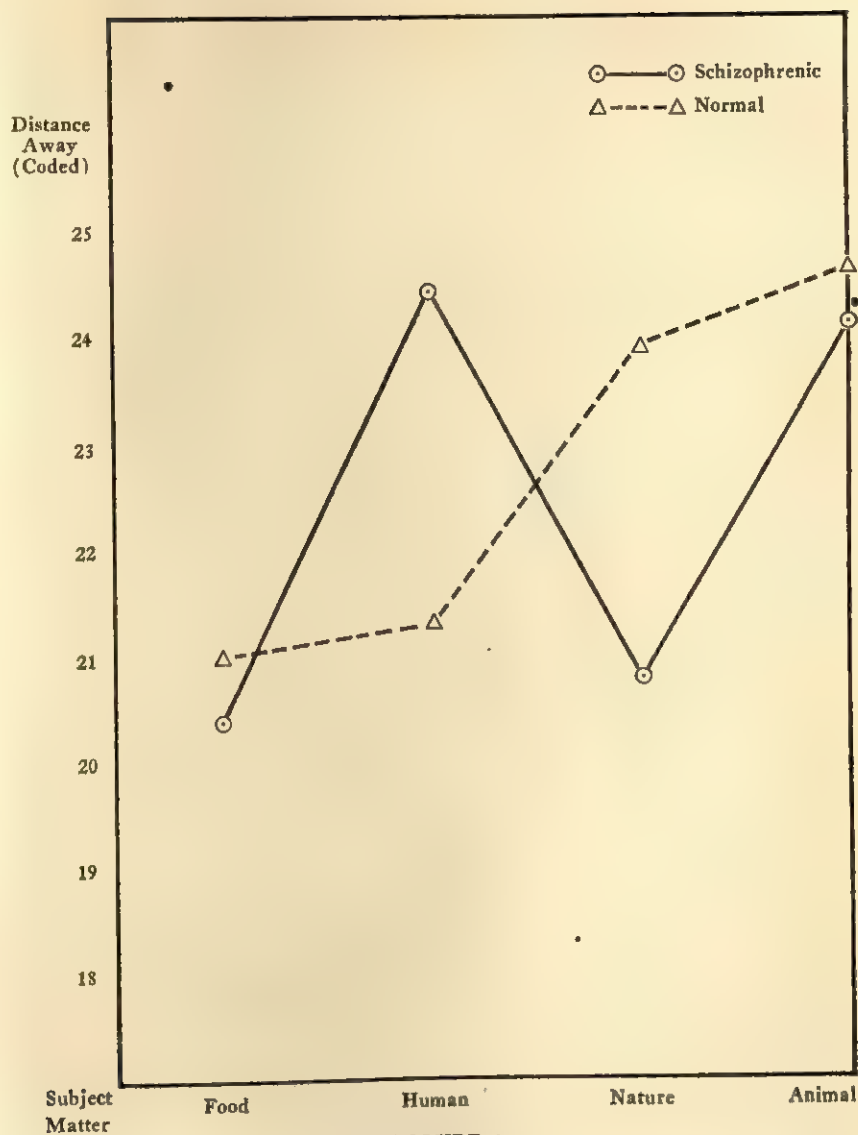


FIGURE 1
INTERACTION BETWEEN GROUPS AND COLOR FOR COMBINED SERIES

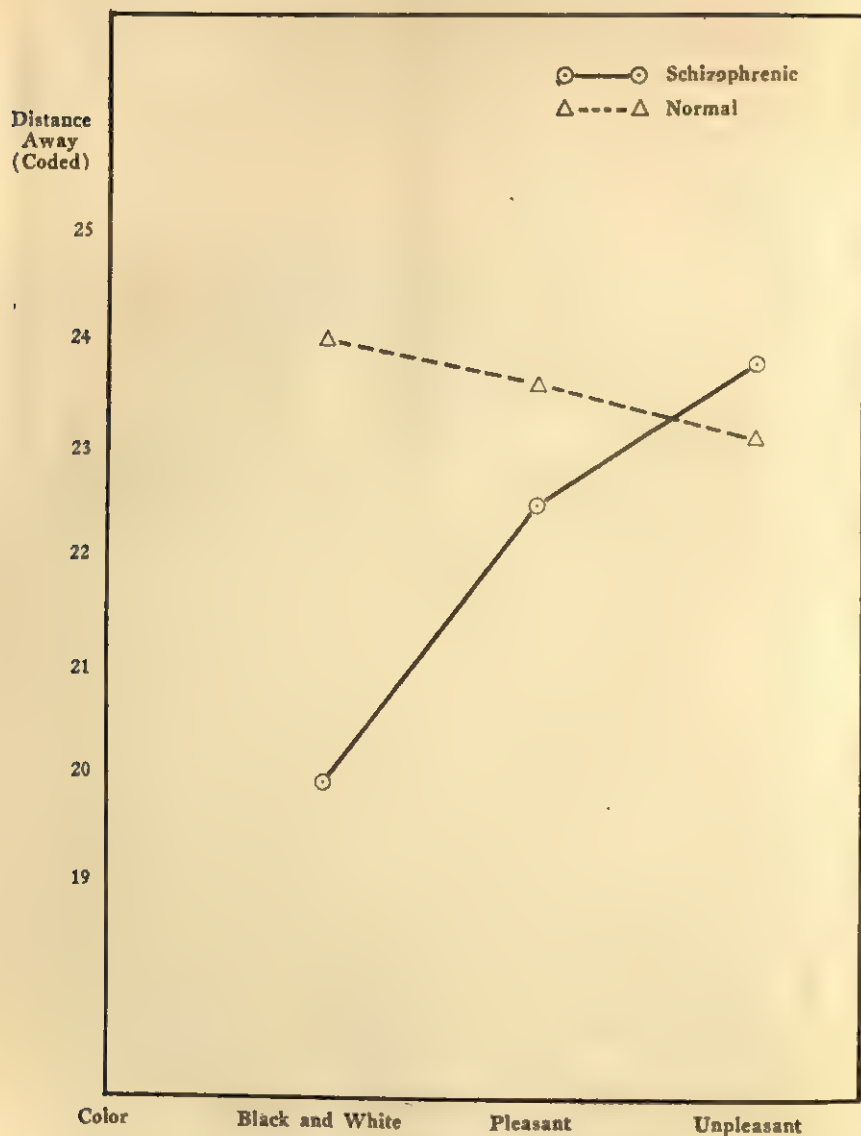


FIGURE 2
INTERACTION BETWEEN GROUPS AND COLOR FOR COMBINED SERIES

D. DISCUSSION

Because, in this research, all subjects placed the images closer than the "required alignment," and Hastings (13) showed that insecure individuals tended to place moveable images significantly closer than did secure individuals, it might appear that all subjects in this experiment were insecure. However, since schizophrenics are far more anxious and less secure than normals and no significant main effect differences between the schizophrenic and normal groups were noted, the level of security probably did not affect image alignment. Rather, because image placement was in all cases closer than required, this suggests that the apparatus successfully reduced veridical distance cues and, therefore, encouraged the use of value and affective variables in image placement in this experiment. Within the framework of closer placement, however, it was noted that the schizophrenics placed human images significantly farther from them than did the normals, a finding consistent with Smith's (27, 28) that insecure subjects placed pictures of human faces farther away than did secure subjects, and with Davis and Harrington's (7) that schizophrenic performance was inferior to that of normals when dealing with human content.

No difference in placement of food and animal images by the two groups was noted. Therefore, the schizophrenics demonstrated no total cognitive-perceptual deficit. In terms of the postulated schizophrenic regression to an early oral stage (10, 11, 26) the lack of significant difference in food image alignment was unexpected, although it was in agreement with Dunn's (9) findings. It seems probable that the food images utilized in the experiment were rewarding, nonthreatening and interesting to all subjects in this study. However, because the schizophrenics placed the rather neutral, non-threatening nature images significantly closer to them than did normals, it may be assumed that normals, being less interested in such neutral stimuli, brought them less close than did the schizophrenics who responded with a willingness to place such images closer to them. Such absence of threat might constitute a reward to a schizophrenic, though this question cannot be answered on the basis of data provided by this research.

However, both groups placed animal images farthest from them, but because one image was of a prowling wolf, it is possible that this image represented a threat to all the subjects who, therefore, responded by placing it farther from them than they did all the other images.

Schachtel (25) considered the relationship between color and affect to be innate and basic, while Phillips and Smith (22) considered reactions to

color an affective experience and Coleman (6) demonstrated the inability of the schizophrenic to respond adequately to affective experiences. The normals in this experiment showed no significant differences in their responses to the achromatic, pleasant and unpleasant chromatic images, which suggests an ability to deal realistically with both neutral and affectively-toned stimuli. Their placing the achromatic and pleasant color images farther away may be indicative of image interest value: the unpleasant colors were unusual and perhaps intriguing, while the achromatic were least intriguing, possibly because their lack of affective tones made them non-rewarding or even mildly unpleasant, for the normals. The schizophrenics, however, probably placed the less affectively toned achromatic images closer to them than the more threatening pleasant and unpleasant color images which demanded an affective response the schizophrenic felt himself incapable of giving. The unpleasant color images were placed farther away than the pleasant color images, indicating they were probably more threatening.

These results, therefore, suggest that schizophrenics and normals respond similarly to nonthreatening and nonaffective response-producing images, but that the schizophrenics place threatening and affective response-producing images farther from them than do the normals.

There are several theoretically possible explanations for these results: learning, abnormal biochemical functioning, or an interaction between these.

Placing images as far away as possible may represent learned withdrawal: unwillingness or inability to respond adequately to the type of stimuli the images picture. If so, the schizophrenic may be displaying a learned withdrawal from traumatic experiences with people upon being faced with pictorial representations of such experiences. They have undergone so many real emotionally unpleasant experiences that they have learned to withdraw or to place themselves at a distance from all unpleasant affective experiences. Weakland (29) theorized that in schizophrenogenic families the schizophrenic is continually faced with doublebind situations in which he receives simultaneously contradictory messages from people in authority over him and upon whom he is dependent. It is important to deal with such messages but impossible because the messages are contradictory and the contradictions and inconsistencies are concealed, denied, or inhibited, usually unconsciously. In such a dynamic setting, a schizophrenic might learn to avoid people who give such messages. Because an adequate affective response to simultaneously contradictory messages is impossible and any response would be inadequate and consequently punished or nonrewarded, the schizophrenic might well learn

to withdraw from all affectively-toned experiences. Winder (31) further pointed out that schizophrenics appear to have experienced unusually deviant interpersonal environments which could easily have led to learning experiences in personal and in affective spheres which differ significantly from the experiences of normal individuals.

Heath (14) showed more response disorganization in schizophrenics when anxiety-producing topics or contents were involved than when neutral stimuli were used, and McReynolds (20) generally assumes anxiety to have ultimately physiological and anatomical bases and to result from any external or internal stimulus which is seen as a threat. The results of this study would, therefore, seem to support the theory of an abnormal breakdown in biochemical functioning in specific areas of the central nervous system under the stress, threat, or anxiety produced by chromatic and by human images, and of a normal biochemical functioning under nonthreat conditions.

It is also possible, of course, that the interaction between learning and biochemical variables in schizophrenia differs from their interaction in normality, but research regarding the interaction, as well as further physiological and anatomical studies in the nature of schizophrenia will have to be undertaken before the problem of schizophrenia becomes truly understandable.

E. SUMMARY

This study was designed to determine whether the responses of schizophrenics to several types and colors of images differed from those of normals in general, or only in relation to specific stimuli, and, if the latter, to determine some of the types of stimuli which elicited the abnormal responses.

Fifty schizophrenics and 50 normals were asked to align projected images of food, people, nature, and animals, with the second of five minimally-lighted plastic rods, both rods and images being viewed monocularly.

The results indicated that there was no difference between the two groups in the placement of food and animal images, but that the schizophrenics placed nature images significantly closer, and human images significantly farther, from them than did the normals. While normals showed no significant difference in their placement of the achromatic, pleasant and unpleasant chromatic images, the schizophrenics placed achromatic images closer to them than chromatic images, and placed pleasant color images closer than unpleasant color images.

Possible theoretical implications of learning, of abnormal biochemical functioning based upon threat or anxiety, or of an interaction between

these were discussed in terms of an indication of the probable line of further research to amplify these significant findings.

REFERENCES

1. ABOOD, L. A chemical approach to the problem of mental disease. In Jackson, D. D. (Ed.), *The Etiology of Schizophrenia*. New York: Basic Books, 1960.
2. AMES, A. Visual perception and the rotating trapezoidal window. *Psychol. Monog.* 1951, **65**, No. 7.
3. BEACH, F. Body chemistry in perception. In Blake, R. & Ramsey, G. (Eds.), *Perception: An Approach to Personality*. New York: Ronald Press, 1951.
4. BLAKE, R., RAMSEY, G., & MORAN, L. Perceptual processes. In Blake, R. & Ramsey, G. (Eds.), *Perception: An Approach to Personality*. New York: Ronald Press, 1951.
5. BRONFENBRENNER, U. Toward an integrated personality theory. In Blake, R., & Ramsey, G. (Eds.), *Perception: An Approach to Personality*. New York: Ronald Press, 1951.
6. COLEMAN, J. *Abnormal Psychology and Modern Life* (2nd ed.) Chicago: Scott Foresman, 1956.
7. DAVIS, R., & HARRINGTON, R. The effect of stimulus class on the problem-solving behavior of schizophrenics and normals. *J. Abn. & Soc. Psychol.*, 1957, **54**, 126-128.
8. DENNIS, W. Cultural and developmental factors. In *Perception: An Approach to Personality*. New York: Ronald Press, 1951.
9. DUNN, W. Visual discrimination of schizophrenic subjects as a function of stimulus meaning. *J. Personal.*, 1954, **23**, 48-64.
10. FEDERN, P. *Ego Psychology and the Psychoses*. New York: Basic Books, 1952.
11. FENICHEL, O. *The Psychoanalytic Theory of Neurosis*. New York: Norton, 1945.
12. GOLDIAMOND, I. Perception. In Bachrach, A. (Ed.), *Experimental Foundations of Clinical Psychology*. New York: Basic Books, 1962.
13. HASTINGS, P. A relationship between visual perception and level of personal security. *J. Abn. & Soc. Psychol.*, 1952, **47**, 552-560.
14. HEATH, D. Individual anxiety thresholds and their effect on intellectual performance. *J. Abn. & Soc. Psychol.*, 1956, **52**, 403-408.
15. HELSON, H. *Theoretical Foundations of Psychology*. New York: Van Nostrand, 1951.
16. HINDLEY, J. Value and frequency as variables influencing perception in schizophrenia. *Dissertation Abst.*, 1958, **19**, 877.
17. IMBODEN, J. Brunswiks' theory of perception. *AMA Arch. Neurol. & Psychiat.*, 1957, **77**, 187-192.
18. ITTELSON, W., & CANTRIL, H. *Perception: A Transactional Approach*. New York: Doubleday, 1954.
19. LOVINGER, E. Perceptual contact with reality in schizophrenia. *J. Abn. & Soc. Psychol.*, 1956, **52**, 87-91.
20. McREYNOLDS, P. Anxiety, perception and schizophrenia. In Jackson, D. (Ed.), *The Etiology of Schizophrenia*. New York: Basic Books, 1960.
21. PHILIP, B. Reversals in the perception of Lissajou Figures by psychotics. *Can. J. Psychol.*, 1953, **7**, 115-125.
22. PHILLIPS, L., & SMITH, J. *Rorschach Interpretation: Advanced Technique*. New York: Grune and Stratton, 1953.
23. PIAGET, J. *The Child's Conception of the World*. New York: Harcourt Brace, 1929.

24. RAUSCH, H. Perceptual constancy in schizophrenia: I. Size constancy. *J. Personal.*, 1952, **21**, 176-187.
25. SCHACHTEL, E. On color and affect. Contributions to an understanding of the Rorschach Test, I. *Psychiatry*, 1943, **6**, 393-409.
26. SECHEHAYE, M. The curative function of symbols in a case of traumatic neurosis with psychotic reactions. In Burton, A. (Ed.), *Psychotherapy of the Psychoses*. New York: Basic Books, 1961.
27. SMITH, G. Personality scores and the personal distance effect. *J. Soc. Psychol.*, 1954, **39**, 57-62.
28. ———. Size-distance settings as indicative of personal adjustment. *J. Soc. Psychol.*, 1954, **40**, 165-172.
29. WEAKLAND, J. The "double-bind" hypothesis of schizophrenia and three-party interaction. In Jackson, D. (Ed.), *The Etiology of Schizophrenia*. New York: Basic Books, 1960.
30. WECHOWICZ, T. Autonomic Activity as Measured by Mecholy Test and Size Constancy in Schizophrenic Patients. *Psychosomat. Med.*, 1958, **20**, 66-71.
31. WINDER, C. Some psychological studies of schizophrenics. In Jackson, D. (Ed.), *The Etiology of Schizophrenia*. New York: Basic Books, 1960.

Department of Psychology
Mills College
Oakland 13, California

SERIAL ANALYSIS OF THREE PROBLEM-SOLVING PROCESSES*¹

Michigan State University

D. M. JOHNSON AND J. W. JENNINGS

A. INTRODUCTION

A review of the literature on thinking (2) suggested three separate problem-solving processes, called preparation, production, and judgment, and research was undertaken to separate preparation from the rest by a serial-exposure procedure (3). The production of solutions has also been timed separately, and predictions about relative time spent on preparation and production have been confirmed (4). The present experiment, by including the final judgment of the solutions produced, is a more ambitious attempt at serial analysis and timing of three sequential problem-solving processes.

Judgment has been studied in some detail as a separate process but not as a part of the problem-solving episode. If one uses the model of the animal exploring the maze, entering the goal-box, and being immediately rewarded, judgment need not be considered. Judgment as a separate process can be ignored also when research is limited to problems with a solution, which, when attained, is immediately rewarding, e.g., anagrams. But there are many complex problems with alternative solutions, such as are now used in the study of originality, and judgment of the alternative solutions to these merits its share of attention.

As in a previous study of reciprocal relations between production and judgment (5), plot-titles were used because the writing of titles for short plots has been found to be a good measure of individual differences in originality (7). Thus, to the Ss of the present experiment, preparation consisted of reading the plot, production consisted of writing titles for it, and judgment consisted of selecting the best of these as the final solution. Since judgment has not often been included in research on problem solving, a special examination of this final step was planned. The aim was to identify the three functional components and time them separately, with a minimum of instrumental

* Received in the Editorial Office on March 4, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ The last report of research done under Grant No. G-15909 from the National Science Foundation.

interference, in order to obtain a differentiated picture of the complete problem-solving enterprise.

The central assumption in the attempt to time three consecutive operations is that the three are different. The research mentioned above supported this assumption by using various manipulations of independent variables to influence the different processes differentially, but an alternative approach is correlational analysis of individual differences. This is usually done with measures of achievement, e.g., number of problems solved correctly, but the present analysis deals with the time of each operation. If the three processes are different, correlations between times of different processes, e.g., between production of titles for one plot and production of titles for another, will be high.

If someone produces five solutions to a problem, will he produce his best one first or last? Previous research (1) under similar conditions, showed that the quality of the titles produced remained constant for 12 minutes, and this relationship can be checked with the data to be obtained under the present conditions. Furthermore, when judgment is included, we can raise the more important question whether *S* will select the best of his titles as his final solution and do justice to his productive ability or pass over his best to choose an inferior production.

Next, we must ask the methodological question whether the instrumental interference necessary for serial analysis reduces *S*'s efficiency. The same task of writing five titles and choosing the best can be performed without serial separation of the three processes. A further step is to have *S* write only one best title. Time required and quality of solutions obtained under these two conditions can be compared with comparable data from the serial-analysis condition.

If judgment is to be treated as the terminal process in the solution of certain problems, it is appropriate to ask whether the principles of judgment that have been demonstrated with standard series of stimuli supplied by *E* can also be demonstrated under the present, more loosely controlled conditions. The titles to plots cannot be measured precisely in advance, but they can be rated later with fair agreement by experienced judges. One principle of judgment states that accuracy increases, or difficulty decreases, with an increase in the difference between standard and comparison stimuli (2). When the *S*s select the best of their own productions, there is no designated standard, but it may be assumed that the critical judgment is between the best and the second best. The hypothesis for the present study, then, is that agreement with expert judges increases with an increase

in the difference between the judges' ratings of the best and the second best titles.

Another well established principle concerning time of judgment (2) can only be demonstrated by a separate experiment in which only two titles are produced. The hypothesis for this experiment is that time of judging between two titles increases as the difference between the ratings of the two decreases.

B. PROCEDURES

In Experiment I preparation, which for this problem consisted of reading a plot, was timed by having *S* control the projection of a plot on a screen about 10 feet in front of him. Production was timed by having *S* write with a pen from a penholder with built-in microswitches. The 5 titles were written on separate sheets previously labeled A, B, C, D, and E, and then judgment was timed by having *S* register his choice by pressing one of a bank of five buttons with corresponding labels.

S was seated at an ordinary desk in a dimly lighted room. He was instructed to read the plot, then, when ready, to take the pen from the penholder and write five titles for it on separate sheets of a small tablet, replace the pen in the holder, and then review his five titles, and press the button with the letter of his best one. *E* turned on the projector, beginning the period of preparation. Withdrawal of the pen from its holder by *S* extinguished the projector, turned on a lamp that illuminated the writing area, stopped the first clock which recorded time of preparation, and started a second clock which timed production. Replacement of the pen in a second opening of the penholder stopped the second clock and started a third clock which timed judgment. Pressing one of the five labeled buttons stopped the third clock.

When titles were presented by *E*, the timing of production was of course omitted; preparation and judgment were timed as above. *S* read the plot, then read five titles placed in front of him, and selected the best of these. In order to equate these titles with titles produced by the *Ss*, each *S* was given titles produced by the preceding *S*.

The two plots used in previous research were used again in order to take advantage of the experience of the expert judges, and a third plot, supplied by J. P. Guilford, was added. Each *S* produced and judged five titles for the plots presented first and third, but only judged titles for the plot in second position. The sequence of processes timed was: Preparation₁, Production₁, Judgment₁; Preparation₂, Judgment₂; Preparation₃, Production₃, Judgment₃. The three plots were rotated so that each plot

appeared in each order equally often. Thus the experiment yielded eight times for each *S*: three for preparation, two for production, and three for judgment. Two sets of five titles each were also obtained, and three first choices of titles. Overall time, including instructions and a short break between plots, was about a half-hour.

The *Ss* were college students recruited from elementary classes in psychology. After three *Ss* were randomly discarded to equalize the frequency of the plots in the three orders, data were available for 42 *Ss*.

Experiment II was the same as Experiment I except that the three processes were not separated and only total time was recorded. The plot was projected on the screen and the writing area was illuminated continuously, hence *S*'s task was somewhat simpler. Five titles were written for each of the three plots, rotated as above. Data were obtained from 30 *Ss*.

Experiment III was similar to Experiment II but with two modifications. *S* was told that he could write as many titles as he pleased but was to hand in only the one best one. Thus the button-pressing operation was unnecessary, and the problem-solving episode was terminated when *S* returned the pen to its holder. Data were obtained from 48 *Ss*.

Experiment IV was concerned only with time of judgment. *S* was instructed to write two titles for each plot and then choose the better of the two. Judgment was timed as in Experiment I, with the exception that only two buttons, labeled A and B, were necessary. The three plots were rotated as above. Data were obtained from 16 *Ss*.

C. RESULTS

Table 1 shows mean times for the eight problem-solving processes by order of occurrence in Experiment I. Obviously the largest fraction of the time was spent on production. The second judgment, judgment of others'

TABLE 1
MEAN TIMES IN SECONDS FOR SEPARATE PROBLEM-SOLVING PROCESSES

| Order | Preparation | Production | Judgment |
|-------|-------------|------------|----------|
| 1 | 61 | 211 | 50 |
| 2 | 34 | | 47 |
| 3 | 49 | 184 | 37 |

titles, was not significantly different from the mean of the first and third when compared by a matched-pairs *t*-test ($t = 0.65$, $p > 0.05$). The second preparation, which consisted of preparing to judge others' titles, is shorter than the mean of the other two, which consisted of preparing to produce

and judge one's own, according to a matched-pairs t -test ($t = 4.63$, $p < 0.005$). Hence preparation is more than reading; or reading is different when done for different purposes.

The matrix of correlations computed from the eight times of Experiment I is shown in Table 2. Clusters corresponding to the three problem-solving processes can be discerned and are outlined for clarity. The rotation of plots and the order effects probably reduce the magnitude of the correlations, but the three preparation times are closely related, as are the two production times and the three judgment times. Most of the other correlations are low. It is clear that correlations between times for like processes are high and correlations between times for unlike processes are low. The chief exception is the correlation of .651 between preparation and judgment of the second plot.

TABLE 2
CORRELATIONS BETWEEN TIMES OF EIGHT OPERATIONS

| | Prep ₁ | Prep ₂ | Prep ₃ | Prod | Prod ₂ | Judg ₁ | Judg ₂ |
|-------------------|-------------------|-------------------|-------------------|------|-------------------|-------------------|-------------------|
| Prep ₂ | .610 | | | | | | |
| Prep ₃ | .698 | .753 | | | | | |
| Prod ₁ | .426 | .438 | .336 | | | | |
| Prod ₃ | .550 | .384 | .260 | .847 | | | |
| Judg ₁ | .526 | .298 | .214 | .375 | .438 | | |
| Judg ₂ | .552 | .651 | .423 | .602 | .508 | .517 | |
| Judg ₃ | .477 | .295 | .187 | .389 | .498 | .835 | .511 |

As a check on these clusters that are obvious to inspection, a factorial analysis of the data of Table 2 was carried out by the Varimax method (6) on the Michigan State Integral Computer. Four factors accounted for most of the variance, with large communalities and a definite structure, as shown in Table 3. Factor A is Judgment, though this factor is not as clear

TABLE 3
FACTOR LOADINGS OF TIME DATA OBTAINED BY SERIAL ANALYSIS OF PROBLEM SOLVING

| Process | Factor | | | | h ² |
|--------------------------|--------|-----|-----|-----|----------------|
| | A | B | C | D | |
| Preparation ₁ | .42 | .78 | .31 | .03 | .88 |
| Preparation ₂ | .08 | .71 | .17 | .59 | .88 |
| Preparation ₃ | .03 | .93 | .10 | .20 | .91 |
| Production ₁ | .14 | .15 | .89 | .31 | .93 |
| Production ₃ | .29 | .19 | .91 | .06 | .95 |
| Judgment ₁ | .92 | .14 | .17 | .15 | .91 |
| Judgment ₂ | .36 | .26 | .34 | .78 | .91 |
| Judgment ₃ | .90 | .09 | .23 | .16 | .90 |

as it might be, because Judgment₂, has a smaller loading than Preparation₁. Factor B is clearly Preparation. Factor C is clearly Production. Factor D is due to Judgment₂, judgment of titles supplied by E, and Preparation₂, which immediately preceded it. Preparation and judgment of the productions of others are somewhat different factorially from preparation and judgment of one's own. With the exception of this weak factor the factorial analysis confirms the identification and separation of the three hypothesized processes.

Since each of the 42 Ss of Experiment I wrote five titles for each of the two plots, 420 titles were obtained. Each title was given a rating for originality, aptness, and cleverness on a five-point scale by two graduate students with considerable experience in such ratings from previous research. The correlation between judges for single titles was .38, and, since the means of the two judges' ratings were used, the Spearman-Brown reliability of .55 is applicable. Table 4 shows that when these titles are grouped by order

TABLE 4
QUALITY OF PLOT TITLES AND FREQUENCY OF FIRST CHOICES BY ORDER OF PRODUCTION

| | Order of production | | | | |
|------------------------|---------------------|------|------|------|------|
| | 1 | 2 | 3 | 4 | 5 |
| Mean rating of quality | 2.29 | 2.62 | 2.50 | 2.60 | 2.61 |
| Frequency of choice | 21 | 17 | 13 | 18 | 15 |

of production, mean ratings are quite similar. Analysis of variance yielded an insignificant *F*. This confirms, for a shorter period of production, the linear relation found by Christensen, Guilford, and Wilson. Table 4 also shows the frequency with which *S* chose his preferred title from each of the five positions. The differences here are also not significant, according to a chi-square test. Hence we can say that under these conditions the Ss are as likely to pick their preferred titles from one position as another.

To determine whether the Ss selected their best titles as preferred solutions the mean rating of the judges was taken as the criterion. In Judgment₁ and Judgment₂ each of the 42 Ss chose the best of the titles that he had just produced, and of these 84 choices 30, or 35 per cent, agreed with the expert judges. If we calculate the chance agreement as 1/5 of 84 or 17, this agreement is significantly above chance ($Z = 3.41$, $p < 0.05$). In Judgment₂, judgment of titles produced by others, the agreement was 14 out of 42 or 33 per cent, which is also significant ($Z = 1.97$, $p < 0.05$). It is interesting to note that these Ss judged their own productions about as accurately as they judged others'.

TABLE 5
MEAN TIME IN SECONDS FOR THREE EXPERIMENTS

| Order | Exp. I | Exp. II | Exp. III |
|-------|--------|---------|----------|
| 1 | 322 | 240 | 171 |
| 3 | 270 | 225 | 148 |

Comparison of the three experiments in respect to total time is shown in Table 5. Only the data of the first and third plots are comparable across experiments, since production was omitted from the second plot. The differences between experiments are significant for both the first and third plots, so there is no doubt that the serial-analysis procedure of Experiment I requires the most time. Production of five titles without the separation of the three processes, as in Experiment II, required less time, and production of one title, as in Experiment III, required even less. The difference between the first plot and the third is significant only for Experiment I, according to matched-pairs *t*-tests, presumably because adjustment to the serial analysis was more complicated and allowed more room for improvement with practice.

The three experiments were also compared in respect to quality of solutions. This comparison included the 84 titles which the 42 Ss of Experiment I chose as their best, the 90 titles which the 30 Ss of Experiment II chose as their best, and the 144 titles produced by the 48 Ss of Experiment III. Although the titles from Experiments I and II had been rated previously, these were included with the others, coded, and randomized, and then all were rated blind in the same context. (Because of a clerical error, 60 of the titles from Experiment II were omitted and had to be rated separately. These ratings did not differ significantly from the other 30, presumably because of the extensive experience of the two expert judges, hence all 90 ratings were treated together.) Mean ratings obtained by these procedures were: 2.87 for Experiment I, 2.59 for Experiment II, and 2.63 for Experiment III. The separation of the three processes by serial analysis may have raised the quality of the titles, but the differences between means are not significant, according to an analysis of variance. We can only say that, although the instrumentation involved in serial analysis decreases speed, it does not decrease quality.

The relation between accuracy of judgment and difference between titles is shown in Table 6, arranged from the data of Experiment I. The means of the expert judges' ratings ranged from 0 to 5 by steps of 0.5, but the largest difference between best and second best titles was 1.5. When the difference was this large, all choices agreed with the judges. When the

TABLE 6
FREQUENCY OF AGREEMENT BETWEEN EXPERT'S RATINGS OF TITLES AND Ss' CHOICES
ARRANGED BY DIFFERENCE BETWEEN RATINGS OF BEST
AND SECOND BEST TITLES

| Difference in ratings | Disagreement | Agreement | Chance agreement |
|--------------------------|--------------|-----------|---------------------|
| 1.5 | | 5 | 1 |
| 1.0 | 1 | 4 | 1 |
| 0.5 | 25 | 11 | 7.2 |
| 0 | 28 | 10 | 7.6 |

difference was 0.5, 11 out of 36 choices agreed with the judges, as compared with chance agreement of 7.2. For Table 6 the chi-square is highly significant, hence it may be said that this principle of judgment, derived from psychophysics, applies also to judgment of one's own productions immediately after production. The same relation was found in judgment of titles produced by others, but with only 42 judgments, the chi-square was not significant.

The relation between time of judgment and difference between titles was examined with the 48 titles obtained from Experiment IV. When the title-differences were grouped in steps of 0.5, mean times for each group were not significantly different. When title-differences were correlated with time of judgment, the correlation was .02. The lack of a relation cannot be attributed to unreliability of the ratings because the interjudge correlation was .735 for single titles.

D. DISCUSSION

The serial exposure and timing of Experiment I slowed the Ss, in comparison with Experiment II, by 45 seconds or more. The perceptual-motor features of the task may account for a small part of this difference, but the larger part must be due to the restriction of overlap between processes in Experiment I. In Experiment II writing could start before reading stopped, and judging could start before writing stopped. But, aside from the extra time, no difficulty was encountered. The task seemed reasonable to college students, they followed instructions readily, and the quality of the solutions was not reduced. It is altogether feasible to study problem solving by using problem-solving processes as the units of analysis, units that are smaller than the whole problem-solving episode but larger than the single response.

The correlations and the factorial analysis show that in respect to individual differences in time the three processes are quite distinct. If the problem-solving episode were an undifferentiated whole, or if the three processes overlapped considerably, the pattern of correlations in Table 2 would not

appear. At the same time the three processes should not be considered independent. Reading a plot could be an independent activity, but the data show that time spent reading depends on the activity to follow. Production of titles could not occur without the preparatory reading. And judgment could not occur until the titles to be judged were written. Furthermore, the agreement of the *Ss'* judgments with those of the expert judges shows that the *Ss'* judgments were related to the plots read during the preparatory period. It is reasonable to conclude that three distinct problem-solving processes have been identified, but that they are interrelated serially in the problem-solving episode toward one end, the one best solution.

Separation of one activity from another does not identify either, so one might ask whether the terminal activity can properly be identified as the activity which is usually called judgment. A rational answer is that the function performed is that usually called judgment, i.e., choice of one of a small number of alternatives. An empirical answer is that one principle of judgment, derived from psychophysical judgments, was confirmed when applied to the terminal activity. Another such principle was not confirmed.

E. SUMMARY

Solution of the plot-title problem can theoretically be analyzed into three problem-solving processes: preparation, which in this case consists of reading the plot, production of several titles, and judgment or choice of the best one. One major and three minor experiments were arranged to check on this analysis and to explore some of the relations suggested by it. Judgment was emphasized because this process has not usually been tied in with the solution of problems.

For the major experiment, *S* read a plot projected on a screen, then took a pen from a holder equipped with microswitches and wrote five titles, then replaced the pen and selected the best title for his final solution. Other experiments were variations of this one. Time spent on each operation was recorded. The titles were rated by expert judges.

Factorial analysis of individual differences disclosed three factors, preparation, production, and judgment, as expected, plus a weak factor involving titles written by others.

When five titles were written for each plot, the best title, according to ratings by expert judges, appeared in each position about as often as in any other. The rate of production in respect to quality, was linear. Similarly, the *Ss'* choices of their preferred titles were in one position as often as another.

The Ss' judgments of their own titles immediately after writing them agreed quite well with the judgments of the experts, as did also their judgments of titles written by others.

The data permitted testing two principles of judgment, derived from research with simpler materials. One, rephrased for present conditions, states that accuracy in choosing the best solution increases with an increase in the difference between the best and the second best. This was confirmed, both for choice of one's own titles and for choice of others' titles. Another principle, relating time of judgment and difference between two titles, was not confirmed.

Two additional experiments showed that the serial analysis slowed the Ss, but did not reduce the quality of the solutions. The methodological conclusion is that the solution of certain problems can profitably be analyzed into functional units shorter than the problem-solving episode but longer than the single response. The inclusion of judgment along with preparation and production yields a broad but differentiated outline of problem solving.

REFERENCES

1. CHRISTENSEN, P. R., GUILFORD, J. P., & WILSON, R. C. Relations of creative responses to working time and instructions. *J. Exper. Psychol.*, 1957, **53**, 82-88.
2. JOHNSON, D. M. *The Psychology of Thought and Judgment*. New York: Harper, 1955.
3. ———. Formulation and reformation of figure-concepts. *Amer. J. Psychol.*, 1961, **74**, 418-424.
4. ———. Serial analysis of verbal analogy problems. *J. Educ. Psychol.*, 1962, **53**, 86-88.
5. JOHNSON, D. M., & ZERBOLIO, D. J. Reciprocal relations between production and judgment. *Amer. J. Psychol.*, in press.
6. KAISER, H. F. The Varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 1958, **23**, 187-200.
7. WILSON, R. C., GUILFORD, J. P., & CHRISTENSEN, P. R. The measurement of individual differences in originality. *Psychol. Bull.*, 1953, **50**, 363-370.

Department of Psychology
Michigan State University
East Lansing, Michigan

CONTEXT EFFECTS IN PRODUCTION AND JUDGMENT*¹

Michigan State University

JOSEPH W. JENNINGS AND DONALD M. JOHNSON

A. INTRODUCTION

Research on judgment is quite extensive (3) but this research has typically been conducted without relation to other intellectual processes, using objects, e.g., weights, supplied by the experimenter. Recently, however, Johnson and Zerbolio (5) studied reciprocal relations between production and judgment and demonstrated, among other things, that subjects who practice production of plot titles improve their judgment of plot titles. Johnson and Jennings (4) treated judgment as the terminal process of the problem-solving episode and showed how this process articulated with the others. The subjects shifted readily from production of plot titles to judgment of these titles, and data on the accuracy of judgment of their own productions conformed to predictions based on judgment of objects supplied by the experimenter. The experiments to be reported here were also planned to show reciprocal relations between production and judgment, specifically by way of the context effect.

The context or series effect is an effect of the series of objects being judged on the organization of the scale of judgment. Under certain conditions the objects being judged constitute the effective context so that the distribution of judgments corresponds in a general sense to the distribution of these objects. For example, when a series of auditory frequencies are judged on a scale of two categories, high and low, the PSE or limen between the two categories approximates the mean of the series. The principle is well established (3) though the terminology varies and investigation of the details continues (2, 8). Extension of this effect to the present topic leads to two hypotheses. (a) When production of objects is followed by judgment of such objects, the objects produced constitute a context within which the objects are judged. Hence the category limen of the scale of judgment will approach the mean of the objects produced. (b) When judgment precedes production, the objects judged constitute a context which influences production. Hence, the mean of the objects produced will approach the mean of the objects

* Received in the Editorial Office on March 4, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ Supported in large part by a grant from the National Science Foundation.

judged. Both hypotheses were tested in Experiment I, only the second in Experiment II.

B. EXPERIMENT I

Angles were used because they are convenient for production and for judgment. To test the first hypothesis two groups of subjects were required to produce angles, then to judge angles. One group PJ-S, produced small angles, then judged a standard series of angles. Another group, PJ-L, produced large angles, then judged the same standard series. If the angles produced act as a context, the subsequent judgments of the standard series by the two groups will be different.

Production consisted of freehand drawing of 12 angles on separate sheets of a 5×8 tablet. In order that all subjects have a plausible and uniform orientation to the task they were instructed to draw the first angle with apex to the right, the next with apex to the left, then up, then down, and so on. The size of the angles was controlled by the experimenter's comments, such as "not so large" and "a little smaller than the last one." Thus the PJ-S group drew acute angles ranging in size from about 5° to 55° and the PJ-L group drew obtuse angles ranging from about 125° to 175° .

The standard series, which both groups judged after production, consisted of eight angles ranging from 20° to 160° by 20° steps. These were drawn on 5×8 cards in india ink, with arms two inches long. The apexes of the angles were drawn in four directions, right, left, up, and down, two angles in each direction. The eight cards were presented in random order, then reversed and presented again. Thus each subject made 16 judgments of the standard series. The instructions were simply to judge each "large" or "small."

To test the second hypothesis two groups of subjects were required to judge angles and then produce angles. The JP-S group judged small angles and the JP-L group judged large angles, then each group was asked to draw some angles. If the context of judgment carries over to production, the angles produced by the two groups will be different.

The two series of 12 angles each, constructed as described above, ranged from 5° to 55° with a mean of 30° for the JP-S group and from 125° to 175° with a mean of 150° for the JP-L group. The instructions were to judge each "large" or "small." Having done this, each subject was asked to draw 12 angles, varying the direction of the apexes as above.

When the subjects were queried about the purpose of the experiment, those who had any thoughts on the matter reported that it seemed to center

around the direction of the apex. Apparently the orienting instructions satisfied the subjects' curiosity about the task.

The results from the test of the first hypothesis were in the form of two-category judgments of the standard series. To describe the scale used by each group in judging the angles the PSE or category limen was computed for each group by the Spearman distribution method (9). These limens were 69.0° for PJ-S and 89.0° for PJ-L. Figure 1 shows these limens in relation to the preceding context for each group.

The most direct test of the significance of this difference is a comparison of the number of angles judged "large" by the subjects in each group. The means were 10.1 for PJ-S and 8.1 for PJ-L, and the difference between means is significant at the .05 level. This difference supports the first hypothesis, that the context of production influences judgment.

The results from the test of the second hypothesis were in the form of angles. These were measured by a protractor, without difficulty, and the means were computed as 42.7° for JP-S and 83.2° for JP-L. The difference between these means is significant at the .0005 level. This supports the second hypothesis, that the context of judgment influences production. Figure 1 shows these means in relation to the preceding judgments of each group.

C. EXPERIMENT II

Experiment II was intended to extend the generality of the findings of Experiment I by testing the second hypothesis on verbal material in a group situation. Phrases describing offenses are convenient for this purpose because they can be judged and also produced by college students. It was expected that the context established by judging a series of offenses would influence production of offenses.

The list of 187 offenses published by McGarvey (7) was the source of the material. This list was given to a standardization group of 10 students with instructions to rate each on a scale of seriousness from zero to 100, using numbers such as 10 to 30 for mild offenses, 40 to 60 for moderately serious offenses, and 70 to 90 for very serious offenses. On the basis of this group's ratings 15 highly offensive phrases were selected for the JP-H group, e.g., selling pork from trichinotic pigs, such that the mean rating was 75. The other list, for the JP-M group, was selected from mildly offensive phrases, e.g., keeping a nickel you find in a telephone booth, such that the mean was 36. It was hoped that later list would have a mean of 25 but this was not possible because of the paucity of phrases at the low end of the scale.

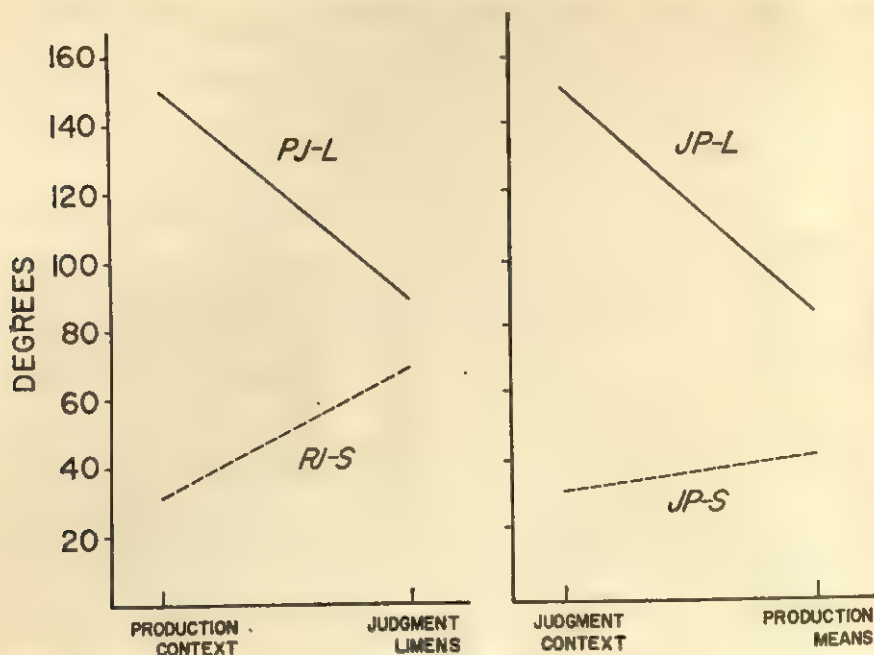


FIGURE 1

EFFECTS OF CONTEXT ON JUDGMENT AND PRODUCTION OF ANGLES

The left graph shows production contexts for two groups and group limens for each group's subsequent judgments of a standard series. The right graph shows judgment contexts for two groups and means of each group's subsequent productions. Both graphs show a significant difference on the second task and also a regression effect.

Two printed forms were prepared from the two lists, with the instructions:

Do this side first. Below is a list of phrases describing offensive acts. Assume the role of an observer of human behavior and rate these phrases so as to indicate how offensive most people would consider them.

Then followed the same instructions on the use of the rating scale as given to the standardization group.

The reverse side of each form was the same for both groups. The instructions read:

We shall need a larger number of phrases to continue our research. Will you write, below, 12 short phrases, each describing a different offensive act.

This was printed on the back of the sheet to minimize the time required for

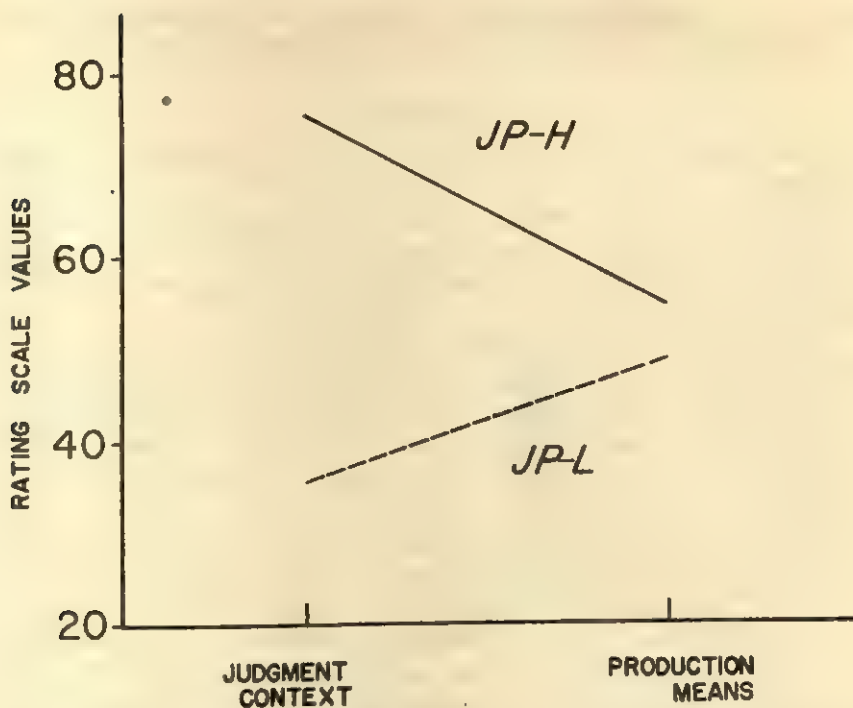


FIGURE 2
EFFECTS OF CONTEXT ON PRODUCTION OF PHRASES DESCRIBING OFFENSES

the subject to proceed from judgment to production and also to remove the printed phrases from view.

The two forms were distributed to an elementary psychology class of 58 students in alternate rows. Thus the JP-H group consisted of 31 subjects and the JP-M group 27.

All phrases produced were coded, typed on separate sheets, and rated "blind" by two graduate students using the scale of zero to 100 described above for the standardization group.² The correlation between the means of the ratings given each subject's productions by the two raters was .79, and since the mean of these two means was used in the statistical analysis, the actual reliability is somewhat above this figure. The group means, calculated from these subject means, were 54.3 for JP-H and 48.0 for JP-M, and the difference between means was significant at the .01 level. Hence the second

² Dominic Zerbolio assisted the first author on this laborious task.

hypothesis is supported with verbal as well as non-verbal material. Figure 2 shows these group means in relation to the preceding judgments of each group.

D. DISCUSSION

Superficially, production and judgment are quite dissimilar, but these experiments have shown that either activity can influence the other. An interpretation in terms of transfer would not be adequate if it assumes that specific responses transfer from one activity to the other but would be adequate if it assumes that the context transfers and that this context controls the specific responses.

Ordinarily, at least when the problem is simple, one produces solutions, then judges them. Hence the PJ sequence is probably the more common one. But in the solution of complex problems, when production and judgment alternate, context effects in the JP direction must be considered. Furthermore one commonly reads and evaluates solutions written by others, then produces solutions under the influence of this context.

Although the expected context effect was significant in all three comparisons, the magnitude and persistence of the effect should not be overestimated. Each graph shows what might be called regression in that the second pair of points is closer to the center of the scale than the first. The most general explanation is that the experimental context has small weight compared to the context of previous experience. If we assume that college students have some previous acquaintance with the whole scale of offenses, from the peccadilloes of undergraduate life to the crimes of the evening paper, such experience will have a regressive influence on the experimental context, lowering the high one and raising the low one. Thus the effect studied here is in a sense a recency effect, since the experimental context is only the most recent of many.

The same explanation holds for the results with angles. In addition, although both experimental contexts were symmetrical around 90° , the lines of both graphs in Figure 1, if extrapolated, would converge considerably below 90° . This is understandable if we assume that most subjects have had more experience with acute than with obtuse angles. Another feature of the PJ sequence is that the standard series of angles presented for judgment also operates as a context, augmenting the regression. In fact this comparison yields a difference only half as large as the JP sequence.

Campbell and others (1, 6) have raised the question whether the context effect in some experiments could be a semantic effect, due to shifts in the choice of words used to define the categories of judgment. In this con-

nection it may be noted that the two JP sequences of the present study have eliminated the semantic interpretation because the second activity is production rather than judgment.

E. SUMMARY

Production and judgment, usually studied separately, were here related reciprocally by an extension of the familiar context effect. One group produced small angles, then judged a standard series of angles. Another group produced large angles, then judged the standard series. The results showed that the scales of judgment used by the two groups differed because of the different contexts established during production. A similar experiment with two other groups, run in the reverse sequence, demonstrated that the context established during judgment of angles influenced the production of angles. And the same effect was demonstrated for judgment and production of short phrases describing offenses. Hence the context effect is not limited to judgment, nor is it a semantic effect dependent on the definition of the categories of judgment.

Although the three comparisons yielded significant differences in the expected directions, the effects are reduced because pre-experimental contexts dilute the effects of the experimental contexts and cause a regression of each group toward the center of each scale.

REFERENCES

1. CAMPBELL, D. T., LEWIS, N. A., & HUNT, W. A. Context effects with judgmental language that is absolute, extensive, and extra-experimentally anchored. *J. Exper. Psychol.*, 1958, **55**, 220-228.
2. HELSON, H. Adaptation level theory. In Koch, S. (Ed.), *Psychology: A Study of a Science*, Vol. 1. New York: McGraw-Hill, 1959.
3. JOHNSON, D. M. The Psychology of Thought and Judgment. New York: Harper, 1955.
4. JOHNSON, D. M., & JENNINGS, J. W. Serial analysis of three problem-solving processes. *J. of Psychol.*, 1963, **56**, 43-52.
5. JOHNSON, D. M., & ZERBOLIO, D. J. Relations between production and judgment of plot titles. *Amer. J. Psychol.*, in press.
6. KRANTZ, D. L., & CAMPBELL, D. T. Separating perceptual and linguistic effects of context shifts upon absolute judgments. *J. Exper. Psychol.*, 1961, **62**, 35-42.
7. MCGARVEY, H. R. Anchoring effects in the absolute judgment of verbal materials. *Arch. Psychol.*, 1943, No. 281.
8. PARDUCCI, A., & MARSHALL, L. M. Effects of mean, midpoint, and median upon adaptation level in judgment. *J. Exper. Psychol.*, 1961, **61**, 261-262.
9. WOODWORTH, R. S., & SCHLOSBERG, H. Experimental Psychology. New York: Holt, 1954.

Department of Psychology
Michigan State University
East Lansing, Michigan



HANDEDNESS AS A VARIABLE OF IMPORTANCE IN DETERMINING APPARENT MOVEMENT DIRECTION*¹

University of Waterloo

MARTIN KRAMPEN

A. INTRODUCTION

A stimulus configuration where one light flash in the center of the screen is followed rapidly by the flashing of two other lights situated equidistantly at both sides of the center light, may be called a symmetric stroboscopic presentation. Exposed to such a presentation, Ss frequently report bilateral apparent movement. The perception of bilateral apparent movement upon viewing a symmetric stroboscopic presentation has been attributed to a "tendency to total assimilation" (7).

It has been hypothesized elsewhere that active practice with one direction in the symmetric pattern will offset this tendency to total assimilation (4, 5). Practice is here defined as the repeated viewing of one-half of the symmetric presentation (e.g., the center light is flashed first and then followed by either the light to the right or to the left). The concept of "active" participation was introduced by Haggard and Rose (2), and, Haggard and Babin (3). In an experiment on the conditioning of perceived direction in autokinetic movement Ss drew heavy arrows (through several layers of carbon paper) pointing in the direction of the perceived movement. "Active" participation then presupposes some reinforcing motor activity during repeated viewing. In the present study, Ss moved with their right hand a lever to the right and pronounced the phrase "to the right" while they were looking at apparent movement to the right.

If a hand movement is involved in active participation, it may be suspected that handedness should constitute a factor interacting with the directional effect of the practice. The increased directional preference acquired through practice should be significantly different for left handed and right handed Ss on the posttest.

* Received in the Editorial Office on March 10, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ This experiment was supported under terms of Research Grant G-8610 from the National Science Foundation entitled "Studies in apparent movement," headed by Dr. Hans H. Toch.

References on handedness and perception are not too frequent in the literature. An interesting though marginal reference on handedness and perception of apparent movement appears in an early apparent movement study of De Silva (1, p. 498):

The question of right-handedness and right-eyedness seemed an interesting cue and two left-handed students were induced to observe. Experimentation with equivocal situations demonstrated a preference for movement in a clockwise direction. In other words, the directional compulsoriness of the movement in the case of the left-handed *O*s was found to be exactly the reverse of that in right-handed *O*s, a discovery which indicates that directional compulsoriness of clockwise movement or of counterclockwise movement may be due to "handedness" or "eyedness."

It was hypothesized in the present study that effects similar to that observed by De Silva should interact with the predicted effect of practice. More specifically, right-handed *S*s should increase from pretest to posttest the number of left preferences, while left-handed *S*s should increase the number of right preferences, although both groups practice with apparent movement to the right.

B. PROCEDURE

1. Method

Viewing apparatus and stimuli have been described in detail elsewhere (4, 6), as has been the *lever apparatus* for active practice (4, 5). *S*s were seated in front of the lever apparatus at about seven feet distance from the viewing box, their head placed in a headrest.

The data from 16 right-handed and five left-handed male and female *subjects* from introductory psychology classes at Michigan State University were evaluated in this study. None of these *S*s had served before in a similar experiment. *S*s who did not see apparent movement, or did not exhibit any "tendency to total assimilation" on the pretest were discarded. *S*s were told that they participated in two separate experiments, one to assess their ability to coordinate a perceptual, motor, and verbal response, the other to find out whether they could spot subtle differences in quality of apparent movement within a symmetric stroboscopic presentation. From the point of view of *S*s, they went first through a "*rehearsal*" of the "first" and the "second experiment." The "*rehearsal*" of the "experiment in coordination" (practice phase) consisted of one lever movement to the right in synchronism with an apparent movement to the right. The "*rehearsal* of the second experiment" was the pretest proper, consisting of nine identical presentations of

the symmetric stroboscopic stimulus pattern. The light in the center of the screen was flashed first at a duration of 250 ms. After an interval of 12 ms., the two lights situated at 1-5/8 inches to both sides of the center light were flashed simultaneously for 250 ms. Instructions hinted at possible changes among the presentations and *Ss* "rehearsed" the spotting of differences in quality between movement to the right and the left of the center.

After the rehearsals the two "real experiments" took place. In the "first experiment" *Ss* received 50 *directional practice* trials, consisting of exposure to an apparent movement to the right, while a lever movement with the right hand to the right had to be executed and the phrase "to the right" had to be pronounced. Since *Ss* were to make three responses at the same time (perceptual, motor, verbal) the representation of the practice phase as "experiment in coordination" was probably convincing to them.

The posttest, perceived by *Ss* as "second experiment" ("spotting differences in quality of apparent movement") was an exact replication of the pretest ("rehearsal of second experiment"). The control of masking pretest, practice, and posttest was deemed necessary in order to exclude the possibility that *Ss* looked through the design of the experiment and reported in compliance with *E's* hypothesis.

In a *posttest interview* *E* tried to assess how "real" *S's* experience of qualitative differences within the symmetrical apparent movement had been. *Ss* were also asked: "Are you by any chance left-handed?" If *S* answered affirmatively, *E* insisted on details as to which activities were performed by *S* with his left hand. A *questionnaire* was given to *S* after the interview. The questionnaire aimed at finding out whether *S* had seen through the experimental design and what his attitude was with respect to the task. Direct observation of *S* during his writing on the questionnaire confirmed whether *S* was left-handed or right-handed.

2. Measures

The raw data of this study were preferential expressions in favour of the left or right direction of apparent movement given after inspection of symmetric stroboscopic presentation (e.g., "The movement was faster to the right"). Reports not expressing any preference were ignored. On pretest and posttest *Ss* had nine occasions to view the symmetric apparent movement stimulus and express their preference.

A pretest score is obtained for each *S* by subtracting the number of left preferences from the number of right preferences. The mean pretest score

for the right-handed group can be compared with the mean pretest score for the left-handed group by *t*-test.

For pretest to posttest change of specific directional preference (e.g., to the right) a change score may be computed: for each *S* a posttest score is computed by subtracting the number of left preferences from the number of right preferences; then the pretest score is subtracted from the posttest score, the difference representing the pretest to posttest change.

With respect to right preference, positive change scores express an increase, negative change scores a decrease, zero change scores no change. By averaging the change scores for right-handed *S*s and for left-handed *S*s separately, two mean change scores are obtained. The significance of the difference between the mean change score of the right-handed and left-handed group can be assessed by *t* test.

C. RESULTS

1. *Pretest to Posttest Change in Right Preferences*

The results of the experiment for right-handed and left-handed *S*s are reported in Table 1.

TABLE 1
COMPARISON OF MEAN RIGHT PREFERENCE SCORES FOR LEFT-HANDED
AND RIGHT-HANDED *S*s

| Groups | <i>N</i> | <i>M</i> | <i>SD</i> | <i>t</i> | <i>df</i> |
|----------|----------|----------|-----------|----------|-----------|
| Pretest | | | | | |
| lf. hds. | 5 | — .200 | 1.720 | .416 | 19 |
| rt. hds. | 16 | .313 | 2.442 | | |
| Change | | | | | |
| lf. hds. | 5 | 4.000 | 3.162 | 2.670* | 19 |
| rt. hds. | 16 | — .813 | 3.395 | | |

* Significant beyond the .05 level of confidence.

Out of 16 right-handed *S*s four increased their number of right preference reports, ten decreased and two did not change. Out of five left-handed *S*s four increased their right preference reports and one did not change. Mean change scores were computed for both groups. The difference between the mean change scores was significant ($t < .05$). The hypothesis of this study that handedness should interact with the directional effect of practice is confirmed.

2. *Right Preferences on the Pretest*

If De Silva's observations would directly apply to the present experiment, a significant difference in right preferences would have to be predicted already on the pretest. Independently from any practice left-handed Ss should then exhibit right preferences and right-handed Ss should exhibit left preferences to begin with. This hypothesis can be tested by comparing the mean right preference score on the pretest for right-handed and left-handed groups. Out of sixteen right-handed Ss, eight exhibited right preference, five left preference and three no preference. Out of five left-handed Ss two exhibited right preference, two left preference and one no preference. The difference between the mean right preference scores is not significant. With respect to right preference on the pretest, left-handed and right-handed Ss seemed to stem from the same population. This might indicate that an effect similar to that observed by De Silva did appear under the present experimental conditions only in interaction with active practice involving manipulation of a lever.

D. DISCUSSION

Table 1 shows that left-handed and right-handed Ss increased the number of directional preferences as a result of practicing a specific direction. However, as predicted by the hypothesis, the directional preferences were significantly different for both groups on the posttest. Practice seemed to sensitize Ss for directional preference in general, while handedness was involved in bringing about the specific directional preference on the posttest. Such a result lends support to the theories stressing the role of past experience (here: stored peripheral cues), and the importance of laterality in perception of apparent movement.

In interpreting the results, the question arises why Ss in this study did not exhibit the effect of handedness on the pretest, if this effect is at all similar to that observed by De Silva. Two main reasons might account for the difference.

First, De Silva used only two left-handed observers, probably for a longer number of observations. Although the N of five left-handed observers in the present study is small, it seems adequate for the use of the small sample statistic t .

Second, the experimental situation was different for De Silva's observers and Ss in the present study. In this respect, one important discrepancy has to do with the stimulus material. De Silva used lines which were related

to each other by a common midpoint but were flashed at different angles, e.g., a vertical line was flashed first and then followed by a horizontal one. Such a presentation could only be resolved perceptually as clockwise or counterclockwise rotation of the vertical line into a horizontal position around a "fulcrum." Any "tendency to total assimilation" is thus prevented by the nature of the stimulus configuration. In the present study three equidistant dots lined up horizontally were used as stimuli. When the dot in the center was flashed first and followed by the simultaneous flashing of the two outside dots, Ss most of the time reported movement "splitting from the center into two directions." Their task was to express a directional preference with respect to the bilateral apparent movement. Only after practice had sensitized them to directionality Ss seemed to become able to use whatever lateral predispositions they had in interpreting the symmetrical apparent movement stimulus preferentially.

E. SUMMARY

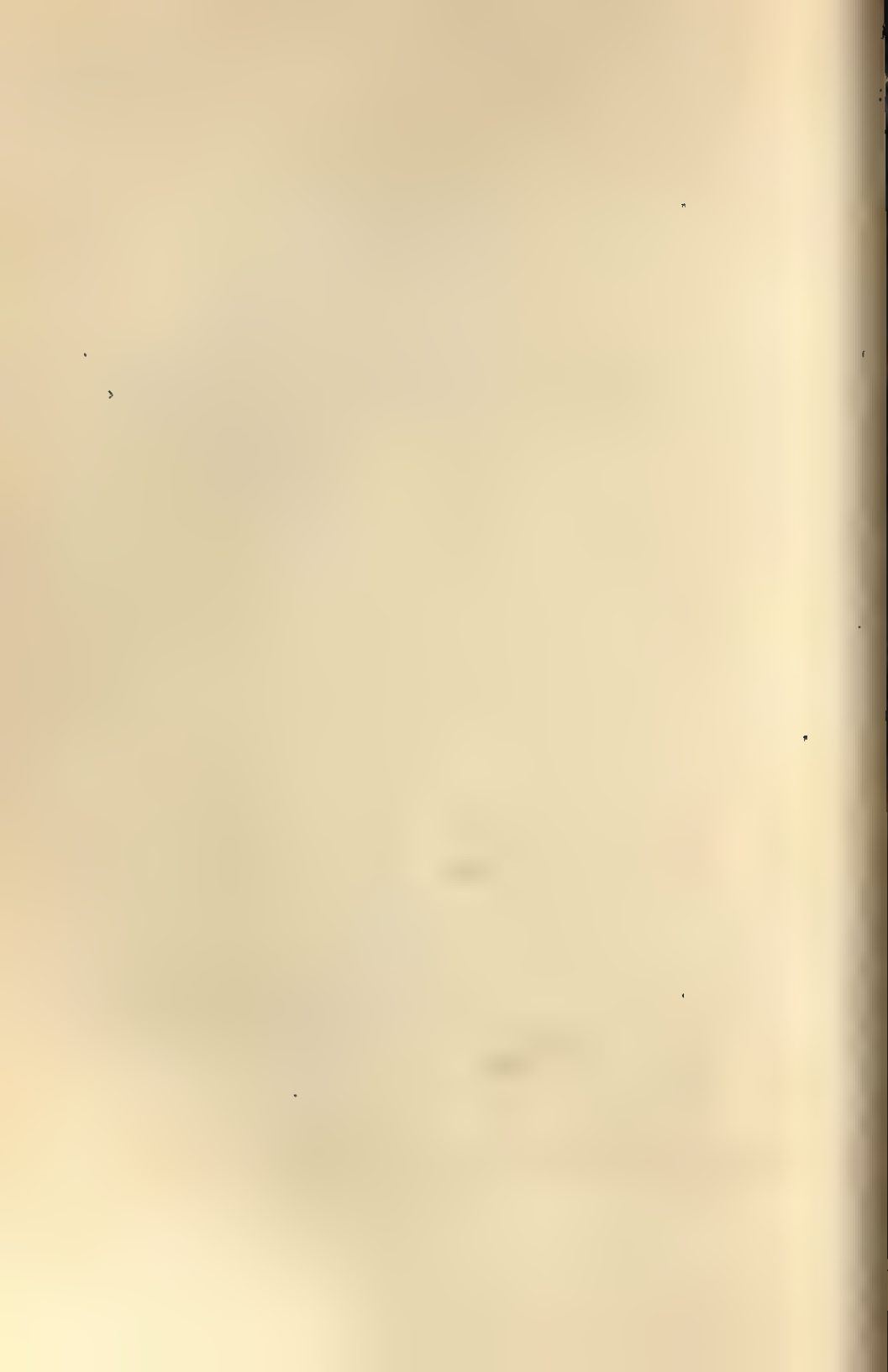
Sixteen right-handed and five left-handed Ss were repeatedly exposed to apparent movement to the right while they actively reinforced this direction by moving a lever with their right hand to the right and saying the phrase "to the right." When viewing a symmetrical stroboscopic presentation after practice, right-handed Ss endowed the left half of that presentation with particular movement quality, while left-handed Ss preferred the right half of the presentation. The result is discussed as lending support to theories stressing the role of past experience (stored peripheral cues) and laterality in perception of apparent movement.

REFERENCES

1. DE SILVA, H. R. An experimental investigation of the determinants of apparent visual movement. *Amer. J. Psychol.*, 1926, **37**, 469-500.
2. HAGGARD, E. A., & ROSE, G. J. Some effects of mental set and active participation in the conditioning of the autokinetic phenomenon. *J. Exper. Psychol.*, 1944, **34**, 45-59.
3. HAGGARD, E. A., & BABIN, R. On the problem of "reinforcement" in conditioning the autokinetic phenomenon. *J. Exper. Psychol.*, 1948, **29**, 511-529.
4. KRAMPEN, M. Some variables of importance in determining reports of directional apparent movement. Unpublished Doctoral dissertation, Michigan State University, 1962.
5. ———. Some variables influencing reports of directional apparent movement. *Percept. & Motor Skills*, 1963, **16**, 349-356.
6. KRAMPEN, M., & TOCH, H. H. The determination of perceived movement direction. *J. of Psychol.*, 1960, **50**, 271-278.

7. VON SCHILLER, P. Stroboskopische Alternativversuche. *Psychol. Forsch.*, 1933, **17**, 179-214.

Design Studies Group
University of Waterloo
Waterloo, Ontario, Canada



"THE MATURITY OF THE CHANCES": A GAMBLER'S FALLACY*

The Pennsylvania State University

WILLIAM M. LEPLEY

A. INTRODUCTION

The writer first encountered the phrase quoted above in an early edition of Hoyle's Games (4). An extended quotation follows: "If the ball in the roulette wheel has not fallen in the red for ten rolls, they think it must come red next time. This is called the 'maturity of the chances' and by betting on this fallacy, many millions have been lost." In the literature of psychology, an early reference to this negative recency effect is found in Thorndike's article (9). With the rise in interest game and decision making theories, numerous related experiments have been reported, for examples, Siegel (8), Jarvick (6), Feldman (2), Myers and Sadler (7) and Edwards (1). With the exception of Siegel, these experimenters have subjected their observers to relatively long and complex programs, apparently, but not likely, assuming that the observer is equipped with some sort of built-in eraser. In Siegel's experiment, each observer was exposed to one condition only, each consisting of a series of forced hand responses, varying from 0 to 160, followed by a free choice series of 10. The experiment described hereinafter resembles that of Siegel, but is judged to approximate more closely a true gambling situation. The experiment was designed to test the reactive inhibition principle in a new context. This principle as it appears in Hull's postulate IX (5), refers to the depression of the probability of a response by successive repetitions of the response.

B. EXPERIMENT

1. Subjects

The subjects for this experiment were 1108 undergraduate students, dealt with in groups varying in size. So far as is known, these subjects were naive with regard to the true purpose of the experiment.

* Received in the Editorial Office on March 15, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

2. Procedure

The several assistant experimenters were provided with appropriate instruction sheets and a fifty-cent piece. The experimental conditions were five in number. In the zero condition, the participants were asked to guess and to record the fall of the coin before the first toss. In condition one, the experimenter tossed the coin and called it Heads, no matter how it fell. The participants were then instructed to call and record the next fall before the toss. In conditions two, three and four, the invitation to guess the next toss was preceded by two, three and four tosses respectively. In each condition all falls were called Heads.

3. Results

The outcomes are shown in Table 1 and Figure 1. The data were halved, and are represented in this form, to show their consistency in a simple

TABLE 1
SUMMARY OF PERCENTAGES GUESSED TAILS UNDER THE FIVE EXPERIMENTAL CONDITIONS
(The data were halved to show consistency)

| Condition | Nt | Half N | % | Nt | Half N | % | Nt | All N | % |
|-----------|----|-----------|------|----|-----------|------|-----|----------|------|
| 0 | 32 | 120 | 26.7 | 26 | 120 | 21.7 | 58 | 240 | 24.2 |
| 1 | 43 | 111 | 38.7 | 35 | 116 | 30.2 | 78 | 227 | 34.1 |
| 2 | 77 | 112 | 68.8 | 80 | 112 | 71.4 | 157 | 224 | 70.1 |
| 3 | 61 | 105 | 58.1 | 62 | 105 | 59.0 | 123 | 210 | 58.6 |
| 4 | 56 | 103 | 54.4 | 59 | 104 | 56.7 | 115 | 207 | 55.6 |

fashion. The values shown are percentages guessed Tails. Inspection of these representations appears to reveal that the reactive inhibition principle holds only through condition two.

4. Discussion

The flexion beyond the condition two point was not wholly unexpected. The limit to five conditions was set on the basis of pilot studies conducted in a similar way. Adding condition six (five preceding calls of Heads) yielded gratuitous recorded comments indicating the onset of suspicion; comments such as: "It stuck to the ceiling." "It stood on edge." "What the Hell goes on here?" Although, there was no verbalized evidence in the records, from the experiment proper, it is judged likely that symbolically represented hypotheses began to emerge to produce the flexion noted.

One incidental observation is appropriate for mention. Goodfellow (3)

reported data showing a 78.2 per cent preference for Heads in first toss guesses (choices). The data from the present study show a similar value; 75.8 per cent. The ratio of the difference to its standard error is $2.40/3.92$ or 0.61 .

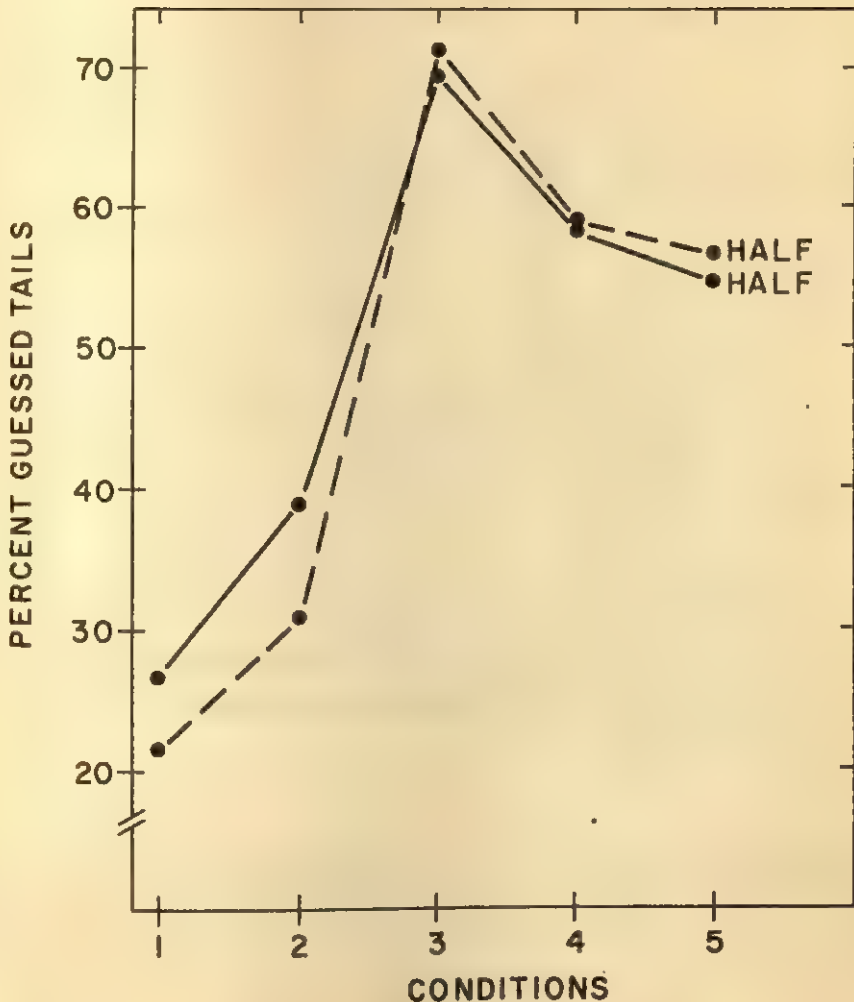


FIGURE 1
REPRESENTATION OF PERCENTAGES GUESSED TAILS UNDER
THE FIVE EXPERIMENTAL CONDITIONS
(The data were halved to show consistency)

C. SUMMARY

One thousand one hundred and eight subjects were divided into five classifications and each group was subjected to a different condition, wherein they were asked to guess the fall of a coin. In the zero condition, they were asked to guess the first toss. In conditions one, two, three and four, their guesses were requested after a varying number of preceding tosses, arbitrarily called Heads; these tosses numbered one, two, three and four respectively. The data appear to show that the reactive inhibition principle held only through condition two (two preceding tosses).

The bias for calling Heads on the first toss appears to have persisted over a period of more than twenty years; and what is more impressive, in approximately the same extent.

REFERENCES

1. EDWARDS, W. Probability learning in 1000 trials. *J. Exper. Psychol.*, 1961, **62**, 385-394.
2. FELDMAN, J. On the negative recency hypothesis in the prediction of a series of binary symbols. *Amer. J. Psychol.*, 1959, **72**, 597-599.
3. GOODFELLOW, L. D. A psychological interpretation of the results of the Zenith Radio experiments in telepathy. *J. Exper. Psychol.*, 1938, **23**, 601-632.
4. HOYLE, E. *Hoyle's Games*. New York: A. L. Burt, 1907.
5. HULL, C. L. *A Behavior System*. New Haven: Yale Univ. Press, 1952.
6. JARVICK, M. E. Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *J. Exper. Psychol.*, 1951, **41**, 291-297.
7. MYERS, J. L., & SADLER, E. Effects of range of payoff as a variable in risk taking. *J. Exper. Psychol.*, 1960, **60**, 306-310.
8. SIEGEL, P. S. Reactive inhibition as a function of number of response evocations. *J. Exper. Psychol.*, 1950, **40**, 604-608.
9. THORNDIKE, E. L. The refractory period in paired associates. *Psychol. Rev.*, 1927, **34**, 234-236.

Department of Psychology

Burrowes Building

The Pennsylvania State University

University Park, Pennsylvania

EARLY FEEDING AND BIRTH DIFFICULTIES
IN CHILDHOOD SCHIZOPHRENIA:
AN EXPLANATORY NOTE*

Psychiatric Institute, Columbus, Ohio

BENJAMIN PASAMANICK AND HILDA KNOBLOCH

In a recent very interesting paper (10) Osterkamp and Sands present provocative data which can probably be interpreted in a more parsimonious manner than they did. They reported:

Histories of schizophrenic and neurotic children were compared as to the occurrence and length of breast feeding and as to the presence of birth and pregnancy difficulties. It was found that the mothers of schizophrenic children were more likely to have attempted breast feeding, but that these attempts were relatively unsuccessful. Birth difficulties also occurred more often in the case of schizophrenic children.

In neurotic children on the other hand there were significantly less pregnancy and birth difficulties as well as less breast feeding. "The results were interpreted in terms of the mothers' unconscious negative feelings towards the infants."

They point out, "The greater frequency of birth complications in the schizophrenic group might be interpreted in terms of the 'continuum of reproductive casualty' hypothesized by Pasamanick, Lilienfeld, and others" (8, 9, 13). They go on to say, however, ". . . since birth difficulties differentiated the schizophrenic and neurotic groups only when associated with short breast feeding, it is likely that the birth difficulties often represent physical manifestations of the mothers' unconscious negative attitudes towards the children."

One of the most important causes of failure to continue breast feeding has been the inability of the infant to secure sufficient milk from the breast. This has been reported by mothers in various ways such as "insufficient milk," "baby gets tired," or "baby goes to sleep." The common denominator to a good many of these breast feeding failures is the exceedingly high negative pressure required to secure milk from the breast compared to the bottle with its adjustable nipple hole. The sucking strength required is quite large.

* Received in the Editorial Office on March 18, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

In addition, the neuromotor integration necessary for this important behavior pattern of sucking and the swallowing which follows is quite complex. It is not surprising, therefore, that this is one of the earliest and commonest patterns to be disorganized in the brain injured infant, and this in turn would lead to abandonment of breast feeding.

In the course of a controlled study of the sequelae of prematurity (5), a number of abnormal behavior patterns were related to the degree of neurologic damage found on examination. The incidence of sucking difficulty during the first months was shown to increase as the degree of brain damage increased. Those infants whose neurologic status is designated in Table 1 as "minor abnormality" comprise the "syndrome of minimal cerebral damage" (4). This syndrome is manifested by minor but clearly defined deviations from the normal neurological and behavioral developmental patterns, usually with more or less complete compensation by 15 to 18 months of age as determined by the standard neurological examination. The differences in the abnormalities from those found in the various types of "cerebral palsy" are

TABLE 1
INCIDENCE OF SUCKING DIFFICULTY DURING FIRST MONTH
RELATED TO NEUROLOGIC STATUS AT 40 WEEKS

| | Neurologic status | | | Total N = 992 | p |
|-------------------------------------|-------------------|-----------------------------|-----------------------------|------------------|------|
| | Normal N = 801 | Minor abnorm. N = 137 | Marked abnorm. N = 54 | | |
| Per cent with sucking difficulty | 11.0 | 16.9 | 27.8 | 12.7 | <.01 |

largely quantitative. The infants exhibiting the syndrome have been shown to have behavioral difficulties later in life.

Social class differences in child rearing practices may also contribute to Osterkamp and Sands' finding that breast feeding of short duration was more common in their schizophrenic children (compared to their neurotics). Breast feeding has been reported to be more common in the lower socioeconomic groups. But we have demonstrated (12) that so are the paranatal abnormalities which are precursors of the continuum of reproductive casualty. The consequent organic brain disease results in early discontinuance of breast feeding because of the sucking difficulty. It is quite striking, however, that in lower socioeconomic groups good prenatal care, and especially protein and vitamin dietary supplements (14), can almost eliminate prematurity, toxemia and abruptio placentae, among the most common abnormalities of pregnancy associated with neuropsychiatric disabilities.

Without using considerable conceptual contortion, it is difficult to conceive of the psychobiologic mechanisms which would produce infectious processes such as pyelitis, or anatomic defects like placenta praevia, also common complications related to prematurity and brain injury. Osterkamp and Sands' explanation, "... that the birth difficulties often represent physical manifestations of the mothers' unconscious negative attitudes towards the children," is hypothetically possible, but good prenatal care has been shown to reduce complications markedly in that group of mothers who could be expected to be most negative to their offspring, that is, girls illegitimately pregnant (11). Since twice as many mothers of neurotic children did not attempt to breast feed their infants at all, it would be just as reasonable to conclude that these mothers had more negative feelings than did the mothers of the schizophrenics.

TABLE 2
PER CENT OF MOTHERS CALLED TENSE RELATED TO NEUROLOGIC STATUS AT 40 WEEKS

| Neurologic status | No. | Mother tense per cent |
|-------------------------|-----|-----------------------|
| Normal | 680 | 2.8 |
| Possible minimal damage | 121 | 4.1 |
| Minimal damage | 137 | 5.8 |
| Possible cerebral palsy | 37 | 10.8 |
| Overt abnormality | 17 | 29.4 |
| Total | 992 | 4.1* |

* $p < .01$.

In the same cohort of prematures and controls mentioned previously, we found that the percentage of mothers called tense increased with the neurologic damage in their infants (Table 2) (5). This is not surprising in view of the fact that the more damaged infants were apt to be cranky, to have other feeding problems in addition to sucking difficulties, to fail to grow well, to be ill more often and to require more hospitalization (7). It is not unreasonable to assume that these anxieties and possible secondary ambivalent or negative feelings might have further disorganizing effects upon the children, thus serving to set off a most unfortunate spiral in mother-child interactions.

It is almost certain that emotional involvement of a negative character might be anticipated in the mothers of "childhood schizophrenics." It is difficult to love a child who does not respond affectively, but it is more likely to be an effect of the behavioral difficulties in the child than a cause.

Another study (6) investigated some 65 preschool children exhibiting

infantile autism, as defined by Kanner (3). When known organic entities, such as phenylketonuria or encephalitis were removed, the remaining 50 children were found to have exceedingly high rates of both complications of pregnancy and prematurity; these were significantly higher than in a control group but similar to those in a group of organically impaired children without autism. These autistic children also had, with only one or two exceptions, neurologic signs of brain damage, ranging from the syndrome of minimal cerebral damage to overt cerebral palsy and mental deficiency. These findings support those of Bender (1), Zitrin (15), and Goldfarb (2), who, in three separate samples of childhood schizophrenics, have found significantly increased rates of prematurity, complications of pregnancy, and definite neurologic findings of cerebral dysfunction. That true schizophrenia in childhood is probably comparatively infrequent is supported by the lack of familial aggregation of schizophrenia in their families as contrasted to the findings in adult schizophrenia and by Goldfarb's demonstration that those children who are obviously organically impaired have relatively healthy parents.

Thus it would seem now that the addition of Osterkamp and Sands' most recent findings to our own and those of others would make it even more advisable at this time to reconsider both the etiology and diagnosis of much of "childhood schizophrenia" or "infantile autism," and possibly assign these cases to the relevant chronic organic brain syndromes.

REFERENCES

1. BENDER, L., & FARETRA, G. Pregnancy and birth histories of children with psychiatric problems. In *Proceedings of the Third World Congress of Psychiatry*, Vol II. Toronto: Univ. of Toronto, McGill Univ. Press, 1962. Pp. 1329-1333.
2. GOLDFARB, W. *Childhood Schizophrenia*. Cambridge: Harvard Univ. Press, 1961. Pp. 216.
3. KANNER, L. *Child Psychiatry*. Springfield, Ill.: Thomas, 1957. Pp. 777.
4. KNOBLOCH, H., & PASAMANICK, B. Syndrome of minimal cerebral damage in infancy. *J. Amer. Med. Assoc.*, 1959, **170**, 1384-1387.
5. ———. The developmental behavioral approach to the neurologic examination in infancy. *Child Devel.*, 1962, **33**, 181-198.
6. ———. Etiologic factors in "early infantile autism" and "childhood schizophrenia." Paper presented at the International Congress of Pediatrics, Lisbon, Portugal, 1962.
7. KNOBLOCH, H., PASAMANICK, B., HARPER, P., & RIDER, R. The effect of prematurity on health and growth. *Amer. J. Public Health*, 1959, **49**, 1164-1173.
8. LILIENFELD, A. M., & PASAMANICK, B. The association of maternal and fetal factors with the development of epilepsy: Abnormalities of the prenatal and paranatal periods. *J. Amer. Med. Assoc.*, 1954, **55**, 719-724.

9. LILIENTFELD, A. M., PASAMANICK, B., & ROGERS, M. E. Relationship between pregnancy experience and the development of certain neuropsychiatric disorders in childhood. *Amer. J. Public Health*, 1955, **45**, 637-643.
10. OSTERKAMP, A., & SANDS, D. J. Early feeding and birth difficulties in childhood schizophrenia: A brief study. *J. Genet. Psychol.*, 1962, **101**, 363-366.
11. PARMELEE, A. H., JR. Prematurity and illegitimacy. *Amer. J. Obstet. & Gynec.*, 1961, **81**, 81-94.
12. PASAMANICK, B., KNOBLOCH, H., & LILIENTFELD, A. M. Socioeconomic status and some precursors of neuropsychiatric disorder. *Amer. J. Orthopsychiat.*, 1956, **26**, 594-601.
13. PASAMANICK, B., & LILIENTFELD, A. M. Maternal and fetal factors in the development of epilepsy. *Neurology*, 1955, **5**, 77-83.
14. TOMPKINS, W. T., & WIEHL, D. G. Maternal and newborn nutrition studies at Philadelphia Lying-in Hospital. Maternal studies. II. Prematurity and maternal nutrition. In Milbank Memorial Fund, *The Promotion of Maternal and Newborn Health*. Papers presented at the 1954 annual conference of the Fund. New York: Milbank, 1954. Pp. 25.
15. ZITRIN, A., *et al.* Prenatal and paranatal factors in psychiatric disorders of children. Presented at the annual meeting of the American Psychiatric Association, Toronto, 1962.

Research Division
Psychiatric Institute
473 W. 12th Avenue
O.S.U. Health Center
Columbus, 10, Ohio

NEED REDUCTION AND REINFORCEMENT: INCENTIVE VALUES OF VARIOUS GOAL OBJECTS AND NOVELTY*

Department of Psychology, University of Utah

BEATA JENCKS AND PAUL B. PORTER

A. INTRODUCTION

The effectiveness of a novel object (i.e., one never before encountered) as a primary reinforcer has been demonstrated in multiple T maze learning of rats (5). In that experiment, novelty was as effective a reinforcer as food, and more effective than mere removal from the goal box.

The present experiment is a further evaluation of the incentive strength of novelty. In it the incentive of novelty has been systematically compared to its own ingredients, taken one at a time, as constant incentives. The experiment consisted of two parts. In the first, a hierarchy of preferences of 10 basic incentives was established. In the second part, each one of these incentives was tested against a sequence of others.

B. METHOD

Ss were 30 female Holtzman albino rats, 34 weeks old. They were caged individually and had Purina Lab Chow checkers and water *ad lib*. Daily handling and a regular schedule of one supplemental feeding per day of a paste of Lab Chow was begun one week before the experiment. Throughout the experiment the supplemental feeding of the paste was continued, with the rats being fed about one hour after they were returned to their home cages from the maze.

The maze was a single unit T, with grey painted walls and a hardware cloth floor and ceiling. Each arm was 27 inches long, four inches wide, and five inches high. Each arm ended on a wooden table, 23 × 9 inches, which was covered by a removable hardware cloth cage. The cage covered the table for all incentives but the "ascending ladder," the "being picked up," and the "dark enclosure." The maze was evenly lighted from above. Guillotine doors prevented retracing from either arm.

The total experiment consisted of 110 trials for each rat. Each was given five trials per day, about half an hour apart. Between trials, Ss were

* Received in the Editorial Office on March 18, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

returned to their home cages. For the first part of the experiment the rats were randomly assigned to a choice between two of the 10 basic incentives (Nos. 1-10, Table 1). No two rats had the same choices, but each incentive appeared six times (three times on the right and three times on the left). Each rat was confronted with its respective choice situation for a total of 50 trials. Preferences for the basic incentives were analyzed for trials 21 to 50 in order to establish their hierarchical sequence. The remaining 35 incentives were presumed equivalent to the basic 10.

TABLE 1
INCENTIVES, LISTED IN SEQUENCE OF THEIR PRESENTATION

| | |
|------------------------------------|-----------------------------------|
| 1. Goal table carpeted with cotton | 24. Leather glove |
| 2. Ascending ladder on goal table | 25. Wet sponge |
| 3. Dish with lab chow paste | 26. Ball of black yarn |
| 4. Being picked up immediately | 27. Crumpled newspaper |
| 5. Fan blowing onto goal table | 28. Ticking alarm clock |
| 6. Old, live male rat | 29. Pencil |
| 7. Perfume stick | 30. Standing mirror |
| 8. Dark enclosure over table | 31. Comb |
| 9. Nails strewn about table | 32. Shoe |
| 10. Shredded paper, 2" deep | 33. Glass plate, covering table |
| 11. Sawdust, 1/2" deep | 34. Rats' droppings |
| 12. Crumpled wire | 35. Loose string |
| 13. Jar lid filled with water | 36. Used, but empty ashtray |
| 14. Cigarette | 37. Rubber bottle stopper |
| 15. Dry dust cloth | 38. Standing empty glass jar |
| 16. Piece of soap | 39. Jackknife |
| 17. Book | 40. Yellow plastic egg |
| 18. Red light bulb | 41. Piece of sheepskin |
| 19. Cat's mandible | 42. Two small pieces of plywood |
| 20. Bunch of keys | 43. Cotton moistened with ammonia |
| 21. Two pieces of chalk | 44. Pebbles |
| 22. Wet washcloth | 45. Open bottle of glue |
| 23. Two flashlight batteries | |

On trials 51 to 60, *Ss* were just picked up at the empty goal tables, so that the effects of the choice situations would be at least partially extinguished.

For the second part of the experiment, the maze had a constant incentive "A" in one goal and a series "B" in the other. The "A" incentives were chosen from the ten basic incentives. "A" incentives were opposed by a series of incentives "B₁, B₂, . . . B_n" in the other goal of the T maze. The "B" incentives were changed as soon as they had been encountered once, and none was used again for any *S*. For trials 61 to 110, half of the rats were randomly assigned to encounter novel objects in the right goal, and the other half in the left goal. Each "A" incentive appeared for three *Ss*. The "B" incentives were used successively in the same order for all rats,

starting with Incentive 1 of Table 1; however, for each individual *S*, the incentives encountered in the first part of the experiment and the one offered as the "A" incentive in the second part were left out of the "B" sequence. Further, the animals were divided into two groups of 15 *S*s each for trials 61 to 110. For one group, Incentive 6 (the rat) was included as a "B" incentive; for the other group it was left out. Choices made on trials 81 to 110 were analyzed to establish the preference rating of the novelty series.

C. RESULTS AND DISCUSSION

The results from both parts of the experiment are summarized in Table 2. Novelty was chosen over all the constant incentives excepting the rat. Further, if the data of rats which encountered this high-valued incentive as either an "A" or a "B" incentive are excluded, novelty is seen to be preferred over all constant incentives.

TABLE 2
RELATIVE FREQUENCY WITH WHICH THE VARIOUS INCENTIVES WERE CHOSEN

| Incentive | Number of <i>S</i> s in Gp | Percentage of choices (of 30/ <i>S</i>) | Percentage of <i>S</i> s choosing above 50% |
|----------------------|-------------------------------|--|---|
| Rat | 6 | 76* | 100* |
| Novelty ^a | 17 | 66* | 88* |
| Novelty | 30 | 65* | 83* |
| Being picked up | 6 | 61 | 83 |
| Shredded paper | 6 | 56 | 83 |
| Perfume | 6 | 50 | 50 |
| Cotton | 6 | 48 | 33 |
| Ladder | 6 | 44 | 33 |
| Fan | 6 | 43 | 17 |
| Food | 6 | 42 | 33 |
| Nails | 6 | 42 | 17 |
| Dark enclosure | 6 | 38 | 0* |

Note: For all incentives but novelty, data are from the first part of the experiment; for novelty, data are from the second part of the experiment.

^a Data from the 17 *S*s for which the rat was neither an "A" nor a "B" incentive.

* $P < .05$ that choice was random.

That a rat would prove to be a strong incentive might have been anticipated had we fully appreciated our own hypothesis—a live and active rat should afford novel stimulation to another, individually caged rat and should do so for a considerable time. Hence, that the rat was more frequently chosen than the novelty sequence may be partially attributable to novelty of the rat.

In general, the rats seemed to like the manipulatable objects. Quite a few nibbled the perfume, the soap, the cigarette, the chalk, and the pencil;

pawed and rooted in the cotton, the shredded paper, and the sawdust; and rolled the nails, the ball of yarn, the batteries, and the plastic egg. Much more manipulative behavior occurred on the novelty side, even when the "A" stimulus was manipulatable (e.g., food, perfume, nails, cotton, or shredded paper). The perfume stick, for example, attracted little attention from any of the nine rats encountering it in the first part of the experiment or as an "A" incentive in the second part of the experiment, but when the perfume appeared only once in the novelty sequence, it was vigorously investigated and was nibbled by 12 of 21 rats. The food seemed to have little incentive value as food and was only rarely nibbled at by two of the three Ss which had it as an "A" incentive. The same food, given later in the home cages, was promptly and invariably eaten.

All animals in the experiment, except one rat which showed a rigid position preference throughout (and, incidentally, against novelty), responded without delay after being put into the starting box. This quick responding occurred in spite of the rather low incentive values of the individual goal objects. Prompt reactivity and general absence of position preference may also be interpreted as indicating the equivalence of the objective incentives and the strong attraction of novelty.

That novelty should stand high in any ranking of incentives is to be expected on a number of counts. A curiosity drive (1) implies a tendency to seek out new stimuli as well as to attend to them. An exploratory drive (8) should lead to investigation of novel stimuli in the immediate environment. Also, as any series of objects will likely include some which are manipulatable, a manipulation drive (7) can be appealed to. On the other hand, neither a need reduction (3) nor a drive reduction (4) hypothesis of reinforcement would predict that animals will seek out novelty or that novelty will have a high value, for it *reduces* neither need nor drive. Novelty-adjustive behavior (6) consists primarily of transforming the novel into the familiar; hence, in a Hullian framework, seeking novelty would seem equivalent to seeking hunger instead of food.

One implication of the present experiment is, we think, that a reinforcing stimulus is any stimulus whatsoever which provokes a response distinct from the response being reinforced. This—obviously Guthrian, possibly Premackian (9)—interpretation is different from a consummation hypothesis (10) as well as from any drive reduction hypothesis. By this interpretation, food, for example, reinforces by provoking eating—a distinctive response; food is a reinforcer because most experimental conditions are contrived to ensure that the food will be so responded to at every presentation. And the same

may be said for all traditional incentives. When the conditions are not so contrived, the traditional incentives may fail to reinforce, i.e., see the literature on drive-incentive consonance, much of which is summarized by Hilgard (2, pp. 211-214). Irrelevant incentives sometimes produce incidental learning, but only on a hit-and-miss basis. An irrelevant incentive has no more impact than an empty alley and no more reinforcing value. The irrelevant incentive or the empty alley, either one, will be a perfectly adequate reinforcer if presented after the animal has been habituated to the rest of the apparatus; provided, of course, that their relative novelty can be maintained over a few trials so that the fact of learning can be assayed.

D. SUMMARY

From choices made in a single unit T maze, preferences of 30 rats were determined for individual incentive objects and for the same objects taken as a novelty sequence. The novelty sequence was found to stand as high as, or higher than, the highest of all but one of those same objects taken individually as constant and familiar incentives. The exception, found to be as enticing as the novelty sequence, was a living rat. These results are interpreted as indicating that all stimuli function as reinforcers simply in proportion to their impact, their effectiveness in eliciting a response.

REFERENCES

1. BERLYNE, D. E. *Conflict, Arousal, and Curiosity*. New York: McGraw-Hill, 1960.
2. HILGARD, E. R. *Theories of Learning* (2nd ed.) New York: Appleton-Century-Crofts, 1956.
3. HULL, C. L. *Principles of Behavior*. New York: Appleton-Century-Crofts, 1943.
4. ———. *A Behavior System*. New Haven: Yale Univ. Press, 1952.
5. JENCKS, B., & PORTER, P. B. Need reduction and primary reinforcement: Varied goal situations as reinforcement in maze learning of rats. *J. of Psychol.*, 1960, **49**, 139-143.
6. McREYNOLDS, P. Exploratory behavior: A theoretical interpretation. *Psychol. Rep.*, 1962, **11**, 311-318.
7. MILES, R. C. Learning in kittens with manipulatory, exploratory, and food incentives. *J. Comp. & Physiol. Psychol.*, 1958, **51**, 39-42.
8. MONTGOMERY, K. C., & SEGALL, M. Discrimination learning based upon the exploratory drive. *J. Comp. & Physiol. Psychol.*, 1955, **48**, 225-228.
9. PREMACK, D. Toward empirical behavior laws: I. Positive reinforcement. *Psychol. Rev.*, 1959, **66**, 219-233.
10. SHEFFIELD, F. D., ROBY, T. B., & CAMPBELL, B. A. Drive reduction versus consummatory behavior as determinants of reinforcement. *J. Comp. & Physiol. Psychol.*, 1954, **47**, 349-354.

Department of Psychology
University of Utah
Salt Lake City 12, Utah

VOTING INTENTIONS AND POLITICAL PERCEPTION*¹

University of California, Los Angeles

HAROLD H. KASSARJIAN

A. INTRODUCTION

Needs, attitudes, values, and other psychological variables are considered by some scholars major determinants of perception. For example, in the field of political behavior and political perception, studies by Berelson, Lazarsfeld and McPhee (1) and McGrath and McGrath (2) indicated that partisan voters tend to perceive their candidate in a favorable light and the candidate's statements and positions taken on issues as favorable to their own point of view, while the opposition is perceived unfavorably.

This study went one step further by trying to determine whether the functional factors of perception are powerful enough to distort a physical characteristic of a presidential candidate. The attempt was made to investigate the relationship between political beliefs—as measured by voting intentions—and the perceived relative height of presidential aspirants.

B. SAMPLE

On October 14, and 15, 1960 the interviewing staff of a marketing research organization conducted face to face interviews in the homes of 3018 registered voters in the State of California. Interviewing on a quota control basis was done in 16 counties selected to represent the various regions of the state.²

By mid-October, 1960, the two major candidates had appeared on the TV Debates three times and hence both had been exposed to the voters in California. Research indicated that 88 per cent of the voters of this state had seen at least one of these programs, in which the two candidates appeared more or less side by side.

* Received in the Editorial Office on March 21, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ A version of this data was presented at the 1962 annual meetings of the Western Psychological Association. Appreciation is expressed to Miss Dorothy D. Corey, Director, Facts Consolidated Division of C-E-I-R, Inc. for permitting the author to collect and analyze this material.

² The results of the poll on voting intentions, when weighted to correct for the sampling error, predicted Kennedy would receive 49.8 per cent of the California vote. The official tally indicated Kennedy did in fact receive 50.1 per cent of the ballots.

C. METHOD

The actual questions posed to the respondents were as follows:

If the presidential elections were being held today, which one of these men would you vote for—Nixon or Kennedy?

Interviewers were instructed to alternate the order of names. If the respondent stated "No opinion" or "Not sure" he was asked:

Which one do you lean towards—Nixon or Kennedy?

Some eight questions later in the same survey, as a rider to the regular commercial poll, respondents were asked:

In your opinion who is taller—Nixon or Kennedy?³

If the respondent answered "Don't know," he was asked:

Would you make a guess please, which candidate do you think is taller?

If still the answer was "Don't know," the interviewers were asked to probe for an answer.

Analysis of the data indicated that there was no significant difference between people who immediately stated their opinion on candidate preference and the persons who were leaning towards a candidate as far as the height variable was concerned. Hence, these two groups could be combined into a single one which had selected a candidate.

However, there was a significant interaction effect on the intensity of opinion as to who is taller. Those persons who guessed that Kennedy is taller rather than "being sure" tended to vote for Nixon. The interaction was not significant among those who guessed that Nixon is taller. Hence, if the "know who is taller" and "guess who is taller" groups are combined the effect is one of making the overall results less significant than they might be. Further analysis indicated this to be true, thus, the two groups were legitimately combined into a single group who "claimed to know" which candidate is taller.

D. RESULTS

Assuming that political beliefs and the selection of a candidate do influence perception we can expect that the people who have no opinion on voting intentions, those who have not yet selected a candidate, will be most unbiased.

³ *This Week* magazine, July 10, 1960, reported Nixon's height as 5'11" and Kennedy's as 6'0".

The results indicated, as can be seen in Table 1, that about half the unbiased population feels that Kennedy is taller and about one-quarter feels that Nixon is taller.

TABLE 1
DATA ON THE HEIGHT VARIABLE

| | Respondent intending to vote for | | |
|-----------------------------------|----------------------------------|---------|------------|
| | Nixon | Kennedy | No opinion |
| Size of sample | 1505 | 1325 | 188 |
| Per cent saying the taller man is | % | % | % |
| Nixon | 42.9 | 23.4 | 29.9 |
| Kennedy | 47.3 | 68.1 | 46.8 |
| Don't know | 9.8 | 8.5 | 23.9 |
| | 100.0% | 100.0% | 100.0% |

Among the Nixon partisans, where the perceptual bias would tend to make Nixon taller, the results indicated approximately a 50-50 split as to which candidate is taller. Among Nixon supporters some 13 per cent more people feel that Nixon is taller than do the "No opinion" voters. A 2×2 chi-square analysis, ignoring those saying "Don't know who is taller" yielded a chi-square of 4.31, significant at the .05 level of confidence. Hence, significantly more Nixon voters than "No opinion" voters feel that Nixon is taller.

Turning to the Kennedy partisans, where the perceptual influence would make Kennedy appear taller, the results indicate that approximately two-thirds of the respondents feel that Kennedy is taller and one-quarter feel that Nixon is taller. Among Kennedy voters some 21 per cent more people feel that Kennedy is taller than among the "No opinion" voters. A chi-square analysis of this data resulted in a statistic of 10.83—significant at the .001 level of confidence. Thus, among the Kennedy voters significantly more individuals believe Kennedy is the taller person than among "No opinion" voters.

The key hypothesis of the study is that there is a positive relationship between political beliefs and the perceived relative height of the candidates—that is, comparing the results of the Nixon and Kennedy voters. A chi square yielded a result of 132.93—by far significant beyond the .001 level, thus confirming the hypothesis. Apparently people who have selected a presidential candidate perceive their candidate as being taller than his opponent.

It might be expected that the more intense the opinion, the greater will

be the distortion effect. As already indicated, two questions were asked with regard to voting intentions; "Whom are you voting for?" and "Whom do you lean towards?" If it can be assumed that those who are leaning towards a candidate hold their opinion less intensely than those who clearly state whom they intend to vote for, there should be less perceptual distortion in the former group, the leaners, than in the voters. As already mentioned, a comparison of these two groups was not significant. Hence, either intensity of opinion is not important, or, and more likely, the questions asked did not in reality measure intensity of opinion. People who said they were leaning towards Nixon were not less partisan than those immediately saying they would vote for Nixon.

A final analysis of the data was carried out to determine whether additional information could be revealed by studying various subgroups. The sample was broken down and analyzed by party affiliation, religion, age, sex and standard of living. In each of the subsamples considered the results were identical to the total sample with chi squares significant beyond the .001 level, except for persons who refused to state their registration or are registered other than Republican or Democrat, in which case the results were not significant.

E. SUMMARY AND CONCLUSION

Several weeks before the 1960 presidential election a representative sample of 3018 California voters were questioned as to their voting intentions and their perception of which of the two major candidates is taller. The major hypothesis of the study was that there is a positive relationship between political beliefs as measured by voting intentions and the perceived relative height of the presidential aspirants. The results indicated that the influence of voting intentions are associated with perceptual distortions not only in the relatively more vague statements and positions taken by the candidate as indicated in previous studies, but go so far as to actually distort a physical characteristic of the candidate—in this case making him appear taller.

REFERENCES

1. BERELSON, B., LAZARSFELD, P. F., & MCPHEE, W. N. *Voting Change*. Chicago: Univ. Chicago Press, 1954.
2. MCGRATH, J. E., & MCGRATH, M. F. Effects of partisanship on perceptions of political figures. *Publ. Opin. Quart.*, 1962, 26, 236-248.

University of California
Graduate School of Business Administration
Los Angeles 24, California

DRUG EFFECTS ON COMPETITIVE-PAIRED ASSOCIATE LEARNING: RELATIONSHIP TO AND IMPLICATIONS FOR THE TAYLOR MANIFEST ANXIETY SCALE*

Massachusetts Mental Health Center

ALBERTO DiMASCIO

A. INTRODUCTION

Recent publications have demonstrated that the administration of psychotropic drugs can alter an *S*'s performance in the learning of a competitive-paired-associate list (CPAL) of words. The theoretical bases for these studies were derived from Hull's and Spence's drive theories and experiments showing the relevance of the role of anxiety in the learning of verbal material in a competitive response situation and from experience accumulated in clinical psychiatry indicating that psychotropic drugs have "anxiety-reducing" properties. The hypothesis followed that if anxiety interfered with an *S*'s learning in a competitive response situation, then drugs that reduced this anxiety should have a facilitative effect on performance in such a task.

Burnstein and Dorfman (2) in their study demonstrated that *Ss*, under the influence of 1200 mg. of meprobamate, a compound cited in clinical reports as capable of decreasing "non-specific anxiety," performed better (fewer trials to criteria) in a CPAL situation than another group of *Ss* receiving a placebo. In their study, Burnstein and Dorfman utilized the Spence (9) CPAL paradigm.

Brown, DiMascio, and Klerman (1) used the Mednick-DeVito (5) CPAL model in their study comparing the effects of a single dose of reserpine (.5 mg.), phenyltoloxamine (400 mg.) and a placebo in the same subjects (normals). They used the Mednick-DeVito paradigm since it was possible with the aid of the Russell-Jenkins (7) norms of the Kent-Rosanoff word list to construct lists of equal difficulty. Self-rating questionnaires and psychiatric evaluations were used to assess the initial level of anxiety and the changes produced by the drugs. It was found that reserpine, an agent that is clinically effective in allaying anxiety, caused *impairment* in the learning of the lists (more trials were needed and a tendency for more errors to be made) in comparison to phenyltoloxamine, a relatively ineffectual drug

* Received in the Editorial Office on March 21, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

clinically. However, in this study at the dosages employed, neither reserpine nor phenyltoloxamine substantially or consistently reduced anxiety, either subjectively or on the psychiatric ratings.

These findings led to the extension and elaboration of studies relating the effects of drugs on the CPAL task. In the last cited study it had been also noted that phenyltoloxamine produced hypnotic effects (similar to the barbiturates) while reserpine did not cause this marked sleepiness. In view of this, the possibility arose that phenyltoloxamine's hypnotic properties rather than its anxiety-reducing qualities may have been the factor responsible for its facilitative effects on the CPAL task.

The present study attempts to clarify the roles of, or the relationship between, the "anxiety-reducing" and "hypnotic" properties of psychotropic drugs and CPAL performance.

B. METHOD AND PROCEDURE

The *Ss* were 16 physically and mentally healthy paid volunteer, male college students. Prior to use in the study, *Ss* were interviewed by a psychiatrist and administered Taylor's Manifest Anxiety Scale.

Each student participated in four experiments carried out at one-week intervals. The *Ss* were told only that this was a drug experiment in which they could receive either a different drug, the same drug, or a placebo on the various experimental days. No mention was made of specific drug names or the class of drugs to be used.

The drugs administered were phenyltoloxamine (100 and 200 mg.), secobarbital (50 and 100 mg.) a hypnotic barbiturate, meprobamate (200 and 400 mg.) and a placebo. Meprobamate, a well-known "anxiety-reducing" drug, was included in view of Burnstein and Dorfman's findings cited above, and also because phenyltoloxamine produced some effects similar to those of meprobamate.

The *Ss* were divided into two groups of eight *Ss*. One group received 100 mg. phenyltoloxamine, 200 mg. meprobamate, 100 mg. secobarbital and a placebo; the other received 200 mg. phenyltoloxamine, 400 mg. meprobamate, 50 mg. secobarbital and a placebo. The *Ss* received either a placebo or a single oral administration of one of the three drugs on four separate experimental days. Drugs were given in identical looking capsules and a double-blind procedure followed. Assignment of drugs was made using a modified factorial design.

The *Ss* were tested on the CPAL task three to four hours after drug administrations (time of peak action of the drugs). A self-rating questionnaire

(3) that included items for assessing changes in "anxiety level" and in the "state of sleepiness," on a seven point scale, was filled out by the *Ss* before the drug was given and just preceding the CPAL testing.

The CPAL task consisted of the learning of lists of 16 pairs of words (by the Method of Anticipation). A Hull-type memory drum was used to present the material. In learning a list, the *S* was first shown all the 16 word pairs once—the stimulus word alone first, and then the stimulus-response pair. The *S* was instructed that when the stimulus word subsequently appeared in the memory drum slot, he was to supply the response word aloud before the stimulus-response pair appeared. All stimulus words and stimulus-response pairs were presented for three seconds each. An *E* seated beside the *Ss* recorded the data (the number of trials to criterion and the number of incorrect responses). A list was considered learned after two consecutive errorless trials of giving the proper response words to the stimulus words. The *Ss* were tested individually.

The CPAL lists were constructed from the Russell-Jenkins norms on the Kent-Rosanoff word list, and the response rates to stimulus words given in the source were treated as probabilities of association.

The competitive response word lists were constructed using Mednick and DeVito's paradigm for a competitive response situation which is,



where *S1* and *R1* have a high probability of association, *S2* and *R2* have a minimal or zero probability of association, and *S2* has a stronger probability of association with *R1* than has *S1*. Thus, a situation exists in which two stimuli may evoke the same response. The *S*'s task is to learn the *S1*—*R1*, *S2*—*R2* etc., pairs.

Since there were four experimental conditions, it was necessary to construct four lists of equal difficulty.

The lists were equated on both the mean probabilities of association and the competition coefficient (1).

Each list contained eight competitive response sets, or 16 word pairs in all, to be learned.

Order of presentation of word pairs in each list was varied to reduce serial position learning effects. Rotation of the lists was counterbalanced with the drug administrations.

C. RESULTS

An analysis of variance (4) indicated that no significant difference existed between the four lists either in trials to criterion or in number of errors; it

was therefore possible to combine the performance measures across the various drug and dosage conditions.

1. Drug Effects

For each group of *Ss* an analysis was made of the CPAL performance under each drug condition with that of the placebo. Table 1 presents the drug-placebo mean differences and the results of the statistical analyses. At the higher doses of meprobamate and phenyltoloxamine, the *Ss* took significantly fewer trials to learn the lists, and tended ($p = .10$) to make fewer errors in the process, than under the placebo conditions. At the lower dose levels the only significant change noted was after phenyltoloxamine under the influence of which the *Ss* took more trials to learn the lists to criteria than they did under the placebo.

TABLE 1
PLACEBO-DRUG MEAN DIFFERENCES IN CPAL PERFORMANCE

| | | Trials | Errors |
|------|-----|-------------------|-------------------|
| Mep | 400 | + .8 | +2.9 |
| | 800 | -2.6 ^b | -5.9 ^a |
| Seco | 50 | -2.7 | -2.6 |
| | 100 | +1.1 | +2.9 |
| PRN | 100 | +2.9 ^b | +5.5 |
| | 200 | -3.9 ^b | -6.7 ^a |

^a $< .10 > .05$.

^b $p = < .05$ Wilcoxon's *T* test (11) for paired replicates.

Note: (-) = fewer trials or errors with drugs than with placebo.

(+) = more trials or errors with drugs than with placebo.

None of the three active agents, at either dose level, produced a marked or consistent alteration in *self-ratings of anxiety*, but individual variations with respect to the direction and degree of anxiety change induced were noted.

Each of the three active drugs induced an increase in *self-ratings of sleepiness*. In each instance the drugs affected more *Ss* to a greater degree at the higher dose level; the actions of phenyltoloxamine being the most pronounced and consistent, those of meprobamate the least.

2. Relationship of TMAS Scores and CPAL Performance

a. Under placebo conditions. Past studies (9, 10), have indicated that performance on the CPAL task is related to an individual's level of anxiety, such that those with low levels of anxiety learn the task more rapidly

and with fewer errors. Accordingly, for the 16 *Ss* the TMAS scores (ranging from three to 24) were correlated with their CPAL performance after the placebo. No significant correlations were obtained between the TMAS scores and both the number of trials to criterion or the number of errors. Under the placebo condition, and *Ss*' initial level of anxiety appeared to have no bearing on CPAL performance.

b. Under drug conditions. The effects of the drugs on CPAL performance measures (as represented by drug-placebo performance differences) were correlated with TMAS scores to determine whether the drugs induced a differential response in individuals varying in initial level of anxiety. Indi-

TABLE 2
CORRELATIONS (RHO) BETWEEN TMAS SCORES AND DRUG ACTIONS
ON CPAL PERFORMANCE

| | | Trials | Errors |
|------|-----|------------------|------------------|
| Mep | 400 | .07 | .24 |
| | 800 | .81 ^a | .65 ^a |
| Seco | 50 | .02 | .55 ^b |
| | 100 | -.12 | -.36 |
| PRN | 100 | .20 | .22 |
| | 200 | .01 | .55 ^b |

^a $p = < .05$ (Rhos, corrected for ties (8, p. 237)).

^b $p = < .10 > .05$.

viduals with higher TMAS scores, under the influence of 800 mg. meprobamate, made fewer errors and took fewer trials to learn the CPAL lists than did *Ss* with lower TMAS scores; there was a similar tendency ($p = .10$) for *Ss* with the higher TMAS scores to learn the lists with fewer errors than the low TMAS scorers after 50 mg. secobarbital and 200 mg. phenyltoloxamine. It thus appears that these drugs, at the doses indicated, produce a more beneficial effect (greater improvement) in *Ss* with higher initial levels of anxiety—as measured by the TMAS.

3. Relationship of Self-Ratings and CPAL Performance

a. Anxiety ratings. As has been previously stated, none of the three active agents produced consistent or marked alterations in self-ratings of anxiety, but individual variations along this dimension were noted. In view of the relationship presented above, correlation coefficients were computed between these drug-induced variations in anxiety ratings, the TMAS scores, and CPAL performance. No statistically significant correlations were obtained. The degree of anxiety change reported after the three active drugs was

independent of the initial anxiety level (TMAS scores) of Ss; it also was not related with performance on the CPAL.

b. Sleepiness ratings. Since the three active drugs induced increases in self-ratings of sleepiness, correlation coefficients were also computed between the "sleepiness" ratings, the TMAS scores, and CPAL performance. The correlations between the TMAS scores and the self-rating changes in sleepiness were computed under each drug condition and the placebo. Ss with the higher TMAS scores reported a greater degree of increased sleepiness than those with lower scores after the placebo, 50 mg. secobarbital and 200 mg. phenyltoloxamine.

TABLE 3
CORRELATIONS (RHO) BETWEEN TMAS SCORES AND SELF-RATINGS OF SLEEPINESS

| | | |
|------|-----------------------|------------------|
| Mep | 400 | .36 |
| | 800 | .41 |
| Seco | 50 | .66 ^a |
| | 100 | .23 |
| PRN | 100 | .37 |
| | 200 | .80 ^a |
| Pl | (average correlation) | .80 ^a |

^a $p = .05$ or less.

CPAL performance was correlated with the changes in self-ratings of sleepiness. The analysis (see Table 4) revealed that those Ss who reported the greatest degree of sleepiness after 800 mg. meprobamate learned the lists

TABLE 4
CORRELATIONS (RHO) BETWEEN CPAL PERFORMANCE AND SELF-RATINGS OF SLEEPINESS

| | | Trials | Errors |
|------|-----|------------------|------------------|
| Mep | 400 | .56 ^b | .23 |
| | 800 | .71 ^a | .64 ^a |
| Seco | 50 | .54 ^b | .11 |
| | 100 | -.04 | -.12 |
| PRN | 100 | .26 | .10 |
| | 200 | .42 | .04 |

^a $p = .05$ or less.

^b $p = .10$.

in fewer trials, and with fewer errors, than those Ss who were either unaffected or were not as sleepy. At 50 mg. secobarbital and at 400 mg. meprobamate the same relationship approached statistical significance for the trials measure. Although phenyltoloxamine produced the greatest increase in the self-ratings of sleepiness, no relationship was found between them and CPAL performance.

D. DISCUSSION

The data indicate that the effects of these psychotropic agents on CPAL performance are a function of the particular drug and dosage utilized. For example, under the influence of 800 mg. meprobamate and 200 mg. phenyltoloxamine in comparison to the placebo condition, the Ss showed a significant reduction in the number of trials necessary to learn the competitive-paired associate lists; further, there was a marked tendency to make fewer errors in the process. After 100 mg. Phenyltoloxamine, however, the Ss took a significantly greater number of trials to learn the lists: there was also a slight decrement in learning after 400 mg. meprobamate.

None of the drugs consistently or significantly altered self-ratings of anxiety,¹ but each of the three psychotropic drugs at the higher doses did produce an increase over the pre-drug level in the self-ratings of sleepiness. The Ss reported the greatest degree of hypnosis after 200 mg. phenyltoloxamine; the next greatest incidence of increased sleepiness after 100 mg. secobarbital.

TMAS scores were not related to CPAL performance under the placebo condition. It was found, however, that the Ss who showed, under 800 mg. meprobamate, the greater degree of improvement in both measures of CPAL performance were those whose TMAS scores were the highest. This relationship was also strongly present between TMAS scores and the "errors" measurement after 50 mg. secobarbital and 200 mg. phenyltoloxamine. Thus, it may be seen that TMAS scores were not related to CPAL performance under the placebo, but there did seem to exist a drug-anxiety interaction in which the drugs produced more of a beneficial effect (improved performance) in high anxiety Ss than in those with low TMAS scores.

According to theory, a high level of drive (supposedly anxiety in this case) interferes with learning in a competitive-paired-associate task; if the drive level is lowered, improved performance is predicted. A phenomenon of this nature (improved learning) appears to have occurred in this study, but no reduction in the level of anxiety (drive) could be demonstrated when the seven point self-rating scale was used as the measure of change in anxiety

¹ Reports from clinical studies indicate that these drugs, particularly meprobamate do reduce anxiety (6). The contradiction between our findings and these are most likely due to their use of chronic administrations of these drugs to populations with pathological levels of anxiety in contrast to our single administration of the drugs to "normal" individuals.

Confirmation of the lack of effect of these drugs, *under the conditions of the study*, on "anxiety level" is available from psychiatric observers ratings on these same Ss. These ratings were obtained during the study but are not reported herein.

level. However, those *Ss* who, under the influence of 800 mg. meprobamate, rated themselves as the most sleepy were also those who showed the greatest degree of improvement in their CPAL performance under the drug. This relationship was also noted with 50 mg. secobarbital and 400 mg. meprobamate. Since an increase in degree of sleepiness can be regarded as a lowering of overall "energy" drive level (mental as well as physical), one may speculate as to whether this "energy" drive is another mechanism, or even the principal mechanism, rather than the "anxiety" drive, involved in CPAL performance changes noted in this study. One may also speculate as to whether the TMAS contains within it items that are directly related to or indicative of one's "energy" level.

Supportive evidence for these contentions is provided by the findings that (a) the TMAS was significantly correlated with the sleepiness rating changes induced by 50 mg. secobarbital and 200 mg. phenyltoloxamine (a similar consistency appeared under the meprobamate conditions); *Ss* who had the higher TMAS scores rated themselves as more sleepy under these conditions; and (b) these *Ss*, in turn, tended to take fewer trials or make fewer errors in learning the competitive-paired-associate lists. These findings indicate the need for further exploration of the Taylor Manifest Anxiety Scale and of CPAL theory for their relationship to "level of hypnosis" as well as "level of anxiety." Nonetheless, it is believed that more discriminating and sensitive means for assessing and quantifying short-term changes in "level of anxiety" must be found. Perhaps more potent psychotropic drugs, which more clearly and markedly alter anxiety levels, should be used which alter one phenomenon (anxiety) without simultaneously altering the other (hypnosis).

The failure to detect any significant drug-induced changes in anxiety level may have been due to the insensitiveness of the seven point self-rating scale used. However, using a similar scale changes in sleepiness ratings were noted after the various drugs.

E. SUMMARY

The effects of a single oral administration of three psychotropic agents (meprobamate, secobarbital, and phenyltoloxamine), each at two dosage levels, and a placebo were tested on a competitive-paired-associated learning task. Self-ratings of "level of hypnosis" (sleepiness) and "level of anxiety" were obtained prior to the administration and at the time of peak drug action. Prior to the actual study, the *Ss* had been administered the Taylor Manifest Anxiety Scale.

CPAL performance under the various drug conditions were examined as

they related to the predrug TMAS scores and to the drug-induced changes in "anxiety" and "sleepiness."

The findings indicated that CPAL performance was a function of the particular drug and dosage condition; and that although CPAL performance was not correlated with the TMAS scores under the placebo condition, the Ss with higher anxiety scores were those who showed the most improvement under certain drug conditions.

The TMAS scores were not correlated with the self-ratings of changes in "anxiety" but were correlated with the self-rating changes in "sleepiness" under certain drug conditions, high anxiety (TMAS) scorers reporting the greatest degree of increased drug-induced sleepiness.

One major implication of these findings was that the TMAS may be measuring some phenomenon or drive other than "anxiety."

REFERENCES

1. BROWN, J. W., DIMASCIO, A., & KLIERMAN, G. L. Exploratory study on the effects of phrenotropic drugs on competitive paired-associate learning. *Psychol. Rep.*, 1958, **4**, 583-589.
2. BURNSTEIN, E., & DORFMAN, D. Some effects of meprobamate on human learning. *J. of Psychol.*, 1959, **47**, 81-86.
3. DIMASCIO, A., HAVENS, L. L., & KLIERMAN, G. L. The psychopharmacology of phenothiazine compounds: A comparative study of the effects of chlorpromazine, promethazine, trifluoperazine, and perphenazine in normal males. *J. Nerv. & Ment. Dis.*, 1963, **136** No. 1, 15-28 (Part I) and 1963, **136** No. 2, 168-186 (Part II).
4. KRUSKAL, W. H., & WALLIS, W. A. Use of ranks in one criterion variance analysis. *J. Amer. Statist. Assoc.*, 1952, **47**, 583-621.
5. MEDNICK, S. A. A learning theory approach to schizophrenia. *Psychol. Bull.*, 1958, **55**, 316-327.
6. REMMEN, E., COHEN, S., DITMAN, K., & FRANTZ, J. R. Psychochemotherapy. Los Angeles: Western Medical Publications, 1962. Pp. 85-86.
7. RUSSELL, W. A., & JENKINS, J. J. The complete Minnesota norms for response to 100 words from the Kent-Rosanoff word association test. Tech. Rep. No. 11, 1954.
8. SIEGEL, S. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw Hill, 1956.
9. SPENCE, K., FARBER, I., & MCFANN, H. The relation of anxiety (drive) level to performance in competition and noncompetition paired-associates learning. *J. Exper. Psychol.*, 1956, **52**, 296-305.
10. TAYLOR, J. A., & CHAPMAN, J. P. Paired-associate learning as related to anxiety. *Amer. J. Psychol.*, 1955, **68**, 671.
11. WILCOXON, F. Some rapid approximate statistical procedures. New York: American Cyanamid Corp., 1949.

Psychopharmacology Research Laboratory
 Massachusetts Mental Health Center
 74 Fenwood Road
 Boston 15, Massachusetts



EFFECTS UPON BRIGHTNESS PRODUCED BY VARYING THE
LENGTH OF THE NULL INTERVAL SEPARATING SUC-
CESSIVE "SINGLE" PULSES: SENSORY IMPLICA-
TIONS OF THE ALTERNATION OF
RESPONSE THEORY, I*¹

Department of Psychology, Michigan State University

THOMAS M. NELSON, S. HOWARD BARTLEY, AND ROBERT M. JEWELL

A. INTRODUCTION

It has long been supposed by Bartley (2) that cortical response amplitude is a neurophysiological correlate of sensory brightness where records are such as to describe the *overall pattern* of function manifested by cell masses in the optic pathway. Such recording is to be contrasted with records obtained from micro-electrode sampling of the activity of various kinds of individual elements (8).

It has long been known (1) that *overall* cortical response to the first stimulus in a train is always larger than those succeeding it. This presumably is due to the fact that a maximal number of channels (the parallel neural units prior to the optic cortex) are ready to respond synchronously.

In contrast, the size of the *overall* response to an identical *second* stimulus seems to depend largely upon two things. First is the length of interval separating the two impingements, that is, the magnitude of the *null period*. If stimuli rapidly succeed each other, the needed period for recovery for units contributing to the total effect is encroached upon and the overall amplitude of cortical response falls. Second, brightness is affected by *autochthonous cortical rhythmicity*. The greatest opportunity to produce a maximal cortical response from successive inputs occurs when stimulus repetition is of the same rate as cortical rhythm. In the adult human this is usually around 10 per second. There are reasons to expect multiples or fractions of this rate to be highly effective also (18). See also discussion in (7).

Maintained stimulation and stimulation with a gradual onset virtually obliterate synchrony. Activity of component elements tends to be distributed

* Received in the Editorial Office on March 27, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ The present paper is a part of the work being done under Research Grant NSF-G-19485 from the National Science Foundation.

randomly throughout because recovery time and latency differ somewhat from element to element.

It follows from what has been said that (a) the magnitude of the brightness response (sensory) to a single short pulse will be greater than that to a prolonged presentation; (b) a null interval of a length adjusting stimulation to autochthonous alpha cortical rhythm will produce an additional brightness effect; and (c) except for effects from autochthonous activity, the magnitude of the brightness response will gradually increase as the null period is extended until, in effect, each pulse is functioning as a single isolated one.

Among the variables that have not been manipulated except incidentally, is length of null period. The foregoing discussion has given some clue as to what to expect were this variable dealt with. It is a study of this sort that is to be reported upon here.

B. METHOD

Pulses were produced by revolving one of two large sectored discs (four feet in diameter) in front of a horizontal slit cut into the front side of a lamp housing. The slit passed diffused radiation from a filament lamp. The one disc was sectored so as to produce a series of pulses 33 msec. long. The other had openings twice this size and therefore produced pulses 66 msec. long. By altering the distance between openings it was possible to progressively increase the null period from 66 msec. to 2366 msec. in the first case and from 33 msec. to 2333 msec. in the second case, while at the same time keeping rate of onset and termination of each pulse constant. Intensity of the intermittent target was pre-set. Two levels of intensity were used, namely 36.3 c/ft² and 3.63 c/ft².

Another identical lamp box was set up along side the first. This target was a steady source of variable intensity. The target radiance was varied by manipulating lamp current and by Wratten neutral density filters. At observation distance the targets subtended 2°5' visual angle in the horizontal direction and 45' visual angle in the vertical. The separation between targets was 36'.

The principal author and two students served as Os. A dot between the two targets was fixated from a chinrest. Ascending and descending matchings (method of limits) of the targets were made and averaged.

C. EVALUATION OF SENSORY RESPONSE

The data are described in terms of a *brightness index* which is a ratio of the mean intensity of brightness matchings for the brief pulses to the mean

intensity of matchings of the same target when continuously illuminated. An index of 100 represents a match just equalling the steady; 50 a match equalling half the intensity of the steady; 200, twice as intense, etc. All points above 100 therefore indicate the occurrence of brightness enhancement or the *Brücke-Bartley effect*, i.e., that the brief photic pulse is brighter than steady illumination of the same intensity. It should be pointed out in passing however, that the index of brightness was devised simply to allow comparisons of effects induced by dissimilar conditions. The indexes are ordinal measurement and not additive in the strict sense. An index of 200 therefore does not necessarily mean that the briefly exposed target looked twice as bright as the steady but rather that the intensity of the steady target had to be doubled to match its brightness. Likewise indexes of 200 do not indicate the same relative increase of brightness when the ranges of intensity matched differ.²

D. RESULTS

Fig. 1 describes the outcome for the principal *O* and combined data from the student *O*s. The data upon which the graph is based are given in Table 1.

TABLE 1
BRIGHTNESS INDEXES CORRESPONDING TO TEMPORAL AND INTENSIVE CONDITIONS USED IN THIS EXPERIMENT

| | | | | | | | | |
|------------------------|----------------------|-------|-------|-------|-------|-------|-------|-------|
| Pulse time 33 msec. | | | | | | | | |
| | PCF | 1/3 | 1/6 | 1/12 | 1/18 | 1/24 | 1/36 | 1/72 |
| | Null period in msec. | 66 | 166 | 366 | 566 | 766 | 1166 | 2366 |
| Intensity | Observer | | | | | | | |
| | N | 139.0 | 74.0 | 100.8 | 138.0 | 180.0 | 212.0 | 201.6 |
| 36.3 c/ft ² | BJ | 149.9 | 97.4 | 76.4 | 75.1 | 95.9 | 110.2 | 142.8 |
| | N | 104.0 | 85.5 | 77.0 | 96.3 | 147.3 | 229.2 | 237.5 |
| 3.6 c/ft ² | BJ | 84.3 | 75.5 | 78.5 | 80.7 | 103.8 | 108.6 | 173.1 |
| Pulse time 66 msec. | | | | | | | | |
| | PCF | 2/3 | 1/3 | 1/6 | 1/9 | 1/12 | 1/18 | 1/36 |
| | Null period in msec. | 33 | 133 | 333 | 533 | 733 | 1133 | 2333 |
| Intensity | Observer | | | | | | | |
| | N | 110.9 | 106.0 | 150.2 | 236.5 | 267.0 | 268.0 | 322.3 |
| 36.3 c/ft ² | BJ | 119.4 | 125.7 | 140.4 | 130.9 | 166.7 | 198.1 | 227.4 |
| | N | 98.3 | 87.9 | 97.9 | 145.0 | 171.7 | 193.1 | 206.8 |
| 3.6 c/ft ² | BJ | 89.9 | 104.6 | 90.6 | 125.5 | 141.5 | 165.4 | 250.2 |

² In Fechnerian psychophysics sensation changes and intensity changes have a logarithmic relation. In Stevens' scheme a power function governs the relation. In no case do the changes bear a simple linear relation to one another.

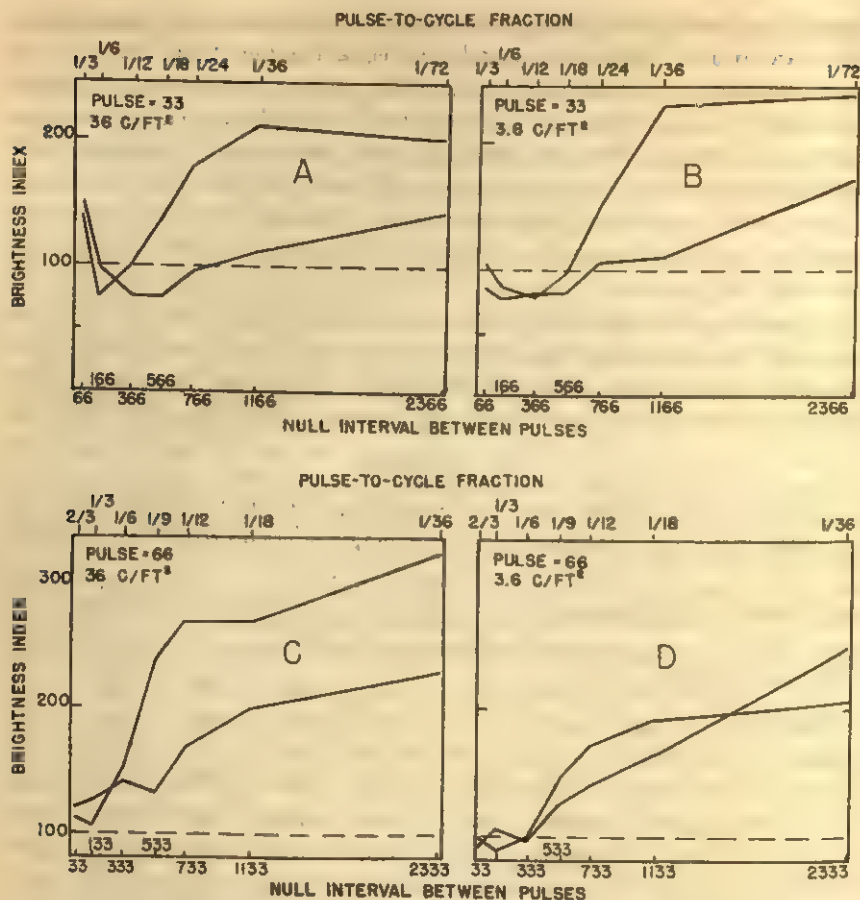


FIGURE 1

BRIGHTNESS INDEXES ASSOCIATED WITH VARIOUS TIME INTERVALS OF NO STIMULATION
 The upper curves are those of the principal investigator, and the lower, the mean values for the student observers.

Much of the sensory outcome suggests operation of a mechanism such as that described by Bartley (2, 3, 6, 8). Individual brief photic pulses when separated by the widest null periods used (2.366 and 2.333 sec.) were seen as brighter than steady illumination of the same intensity in every case. The brightening was almost always considerable in magnitude, the brightness index reaching a maximum of 322. In all cases, brightness gradually declined as the null interval was reduced. In most cases, a null period value was finally obtained in which the brightness of the brief pulse was less than that of

steady illumination. This was not explicitly predicted. But since reducing null period duration (while keeping pulse duration constant) has the effect of increasing rate of stimulation, a point will be finally reached where averaging of illumination will commence to be a factor. The fact that declines from one condition to another were not equally steep or did not drop below the level of the steady at the same null value seems to reflect the fact that the value of critical flicker frequency (CFF), and hence the cycles per second (cps) producing *Talbot effect*, varies with both intensity and PCF (14, 15, 16, 17, 20, 21, 23).

In some cases a sharp rise in brightness occurred where minimal null values occur. For example, when intensity is maximum and the pulse 33 msec., the curves reverse their downward trend and again show brightness enhancement. Notice that the minimum null value is 66 msec. at this point and therefore the pulse-to-cycle fraction (PCF) is .33 and the cps 10. Such a condition produces brief fairly intense stimulation at approximately alpha rate. This is the rate producing peripheral discharges most likely to be favorably received by the cortex. The enhancement indexes are not so large at this point as for more isolated pulses however, indicating that brightness effects produced by maximal utilization of intermittent input are not so great as effects from simultaneous activation of all elements available in the optic pathway. It is with isolated pulses that this availability is expected to be maximized.

There are several ways to manipulate null period (separation between pulses), PCF, and intermittency rate simultaneously while holding intensity constant. Since in this paper we are asking questions primarily about the effects of manipulating null period, let us consider what actually happens to the brightness index when the primary manipulation is duration of null period. The expectations in this case are shown in Figure 2.

Beginning with an indefinitely long null period, one expects a maximal brightness index, because here the pulses are so widely separated in time as to function as single isolated pulses. It is already known that such impingements produce maximal amplitudes of cortical response, and maximal brightness (8). As null period is shortened, various effects occur. Since the sequence depends somewhat upon pulse intensity, we shall consider a fairly high intensity, and a low one, separately. The two parts A and B of the middle component of the overall curve show the expected results for the high and low intensity, respectively.

As null period is shortened from its original length, the resulting bright-

ness index begins to drop. During this shortening, the two other impingement factors of PCF and intermittency rate are also being varied incidentally. Owing to increase in intermittency rate, there come a point at which it reaches the rate of cortical periodicity and this factor would tend to produce brightness enhancement. Hence there would be a reversal of the brightness index trend already in operation, and the brightness index would rise. Thus the curve now goes up, as illustrated in A. With still further reductions in null period, PCF, which has been increasing all this time, now becomes crucial and, as is known (12, 22), reduces brightness. Decreasing null period has the effect of increasing intermittency rate and PCF simultaneously. These changes both work toward producing fusion (9, 10, 11) and hence Talbot level phenomena. As null period becomes still shorter, the condition existing with steady stimulation is approached and therefore the brightness index produced with steady radiance is approached. Theoretically, since null periods have never been obliterated, the brightness index for steady impingement is never quite reached (steady stimulation = PCF of 1.00). Possibly, for all *practical purposes*, it is reached.

Low intensities never produce brightness enhancement (5). Since this is the case, the curve B, as null period is shortened, never should manifest the upward swing indicating brightness enhancement. As the null period is being shortened and at the same time intermittency rate is increasing, fusion (steady brightness) is reached. Talbot levels are now to be considered. These, of course, are brightness levels dependent upon PCF. Small PCF's would produce the low portion of the curve B and as null period is shortened, and PCF is consequently lengthened, the curve should slowly rise, ultimately reaching the level indicating the brightness produced by a virtually continuous radiant impingement. Thus the two different intensity levels just used end by producing similar brightness indexes, *i.e.*, indexes of nearly 100.

Another phenomenon which may have an effect on brightness has not been included in the above description. Bartley (3, 4), Bartley and Wilkinson (13), and Mackavey (19) have studied the mutual effects of the photopic and scotopic systems, which do not always work together temporally and thus at times produce summed brightness, and sometimes discharge at separate instants tending to produce two sensory effects (two flashes) instead of one.

This may produce one of two effects on the early part of the curve in Figure 2 (affects not shown). If the photopic and scotopic discharges were simultaneous, the effect would be a brightness greater than otherwise; if

they occurred in succession, the net effect would be to reduce brightness. Under mid-conditions of intensity and pulse duration, these effects are known not to be simultaneous (3, 4, 13, 19).

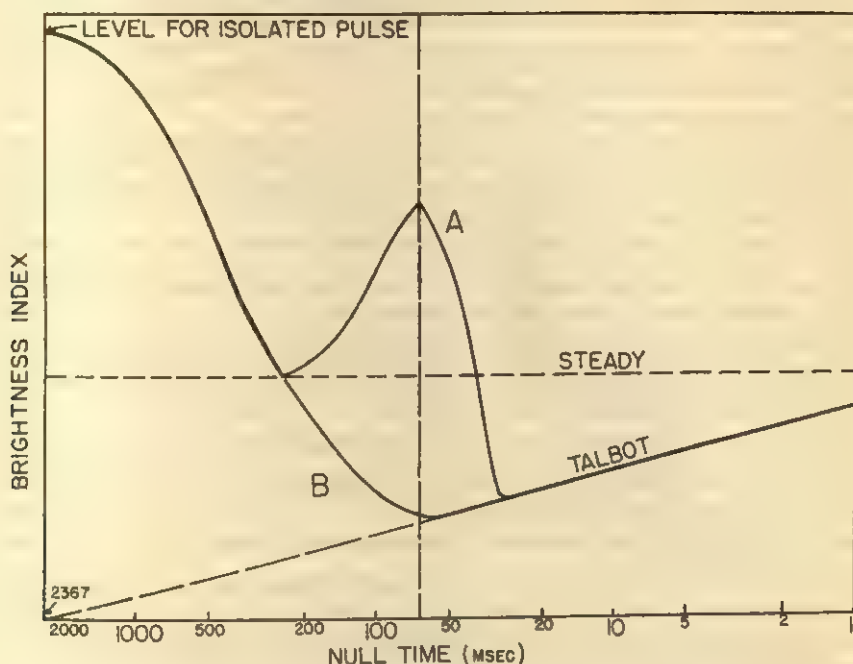


FIGURE 2

A SCHEMATIC PORTRAYAL OF THE RELATIONSHIP BETWEEN THE BRIGHTNESS INDEX AND TIME INTERVAL OF NO STIMULATION FOR A PULSE OF 33 MSEC. AND MODERATE INTENSITY

"A" portion of the curve shows effects for higher intensity inputs and "B" those for lower intensity inputs within a critical range of intervals of no stimulation. The Talbot level for fused targets having the cycle characteristics implied by pulse length and interval of no stimulation is given as a broken line until the point at which the curves intersect is reached.

REFERENCES

1. BARTLEY, S. H. Temporal and spatial summation of extrinsic impulses with the intrinsic activity of the cortex. *J. Cell. & Comp. Physiol.*, 1936, **8**, 41-62.
2. ———. A central mechanism in brightness discrimination. *Proc. Soc. Exp. Biol. & Med.*, 1938, **38**, 535-536.
3. ———. Visual sensation and its dependence on the neurophysiology of the optic pathway. *Biol. Symp.*, 1942, **7**, 87-106.
4. ———. The features of the optic-nerve discharge underlying recurrent vision. *J. Exper. Psychol.*, 1942, **30**, 125-135.

5. ———. Intermittent photic stimulation at marginal intensity levels. *J. of Psychol.*, 1951, **32**, 217-223.
6. ———. Psychophysiology of vision. In Stevens, S. S. (Ed.), *Handbook of Experimental Psychology*. New York: Wiley, 1951.
7. ———. Some facts and concepts regarding the neurophysiology of the optic pathway. *A.M.A. Arch. Ophthalmol.*, 1958, **60**, part II, 775-791.
8. ———. A clarification of some of the procedures and concepts involved in dealing with the optic pathway. In Jung, R. & Kornhuber, H. (Eds.), *The Visual System: Neurophysiology and Psychophysics*. Heidelberg, Germany: Springer-Verlag, 1961.
9. BARTLEY, S. H., & NELSON, T. M. Some relations between pulse-to-cycle fraction and critical flicker frequency. *Percept. & Motor Skills*, 1960, **10**, 3-8.
10. ———. Equivalence of various pulse-to-cycle fractions in producing critical flicker frequency. *J. Opt. Soc. Amer.*, 1960, **50**, 241-244.
11. ———. A further study of pulse-to-cycle fraction and critical flicker frequency. A decisive theoretical test. *J. Opt. Soc. Amer.*, 1961, **50**, 41-45.
12. BARTLEY, S. H., PACZEWITZ, G., & VALSI, E. Brightness enhancement and the stimulus cycle. *J. of Psychol.*, 1957, **43**, 187-192.
13. BARTLEY, S. H., & WILKINSON, F. R. Certain factors in producing complexity of response to a single pulse of light. *J. of Psychol.*, 1953, **35**, 299-306.
14. BROWN, H. C. Some conditions of visual flicker. M. A. thesis, Columbia University, New York, 1929.
15. COBB, P. W. The dependence of flicker on the dark-light ratio of the stimulus cycle. *J. Opt. Soc. Amer.*, 1934, **24**, 107-113.
16. CROZIER, W. J., & WOLF, E. Theory and measurement of visual mechanisms. *V. J. Gen. Physiol.*, 1941, **24**, 635-654.
17. IVES, H. E. Critical frequency relations in scotopic vision. *J. Opt. Soc. Amer.*, 1922, **6**, 254-268.
18. JASPER, H. H. Rapports et communications. *XI Congr. Internat. Psychol.*, 1937, 226.
19. MACKAVEY, W. R. A further investigation of the double flash phenomenon. Unpublished Doctoral dissertation, Michigan State University, East Lansing, 1959.
20. PIÉRON, H. Influence du rapport des phases sur la durée d'interruption d'une stimulation lumineuse périodique à la limite du papillotement. *Comp. Rend. Soc. Biol.*, 1928, **99**, 398-400.
21. ROSS, R. T. The fusion frequency and the light-dark ratio for constant illumination at fusion. *J. Gen. Psychol.*, 1938, **13**, 111-122.
22. VALSI, E., BARTLEY, S. H., & BOURASSA, C. Further manipulation of brightness enhancement. *J. of Psychol.*, 1959, **48**, 47-55.
23. WINCHELL, P., & SIMONSON, E. Effect of the light-dark ratio on the fusion frequency of flicker. *J. Appl. Physiol.*, 1951, **4**, 188-192.

Department of Psychology
Michigan State University
East Lansing, Michigan

BRIGHTNESS CHANGES ASSOCIATED WITH VARIATION
IN THE NUMBER OF PULSES IN A REPETITIVE
PHOTIC TRAIN: SENSORY IMPLICATIONS
OF THE ALTERNATION OF RESPONSE
THEORY, II*¹

Department of Psychology, Michigan State University

THOMAS M. NELSON, S. HOWARD BARTLEY, AND R. M. JEWELL

A. INTRODUCTION

When *mass* recording is made at the cortical level using the rabbit, amplitude of cortical response is greater to the first impingement in a train of physically equal impingements than it is to any succeeding it (2). Bartley theorizes that unequal size of discharge occurs because the number of channels (the parallel neural units discharging into the the optic cortex) ready to respond synchronously (give a summed response) varies with the nearness of the preceding input (3). When photic stimuli are close together in time, full recovery of all channels may not occur and the magnitude of discharge into the optic cortex is thereby reduced.

Figure 1 is a schematic portrayal of a typical cortical record occurring when the stub of the optic nerve of the rabbit is repeatedly stimulated at a rate somewhat above the intrinsic rhythm of the cortex (about five per sec.). Notice that peak response occurs to the first input (A_1). After the first peak, an irregular pattern of response occurs (A_2) although a decline in amplitude occurs as the number of repetitive stimulations is extended. This irregular response section (A_2) Bartley has called a "reorganization period" (1). The reorganization period is stated to be the time it takes for initially variable activity found in the cortex to become modified to permit a regular response to the repeated stimulation (1, p. 48). The reorganization period is terminated when a regular discharge of comparatively low amplitude occurs (A_3). The dotted line fits a curve to the response diagram as a total.

Bartley has hypothesized that amplitude of cortical discharge is positively

* Received in the Editorial Office on March 27, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ The present paper is a part of the work being done under Research Grant NSF-G-19485 from the National Science Foundation.

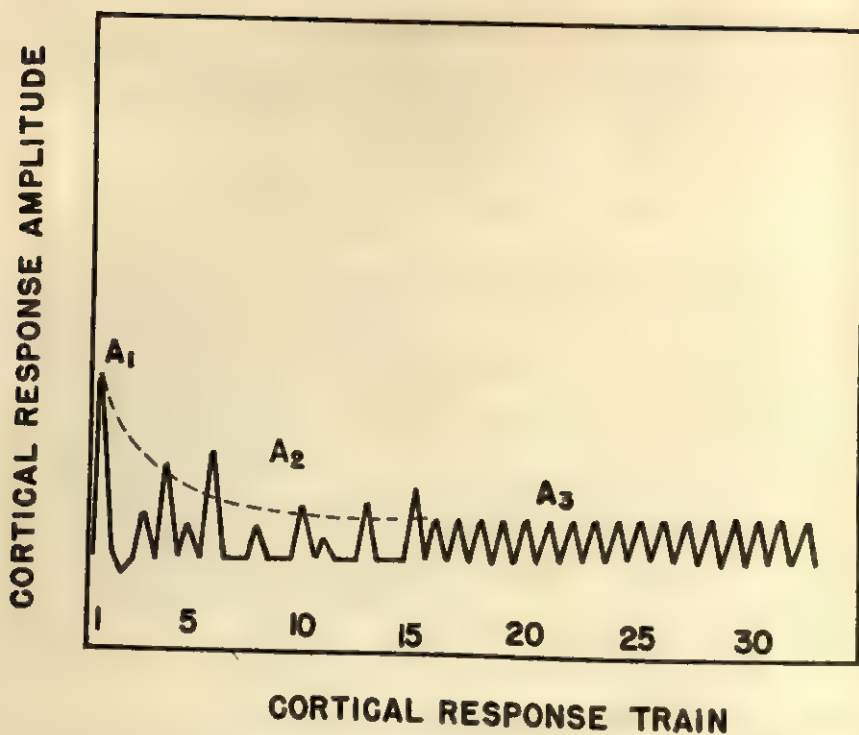


FIGURE 1

SCHEMA OF A TYPICAL *Mass Activity* RECORDING OBTAINED AT THE CORTEX FROM ELECTRICAL STIMULATION TO THE STUMP OF THE OPTIC NERVE OF THE RABBIT AT RATES ABOVE THE INTRINSIC RHYTHM OF THE CORTEX (5 PER SECOND)

Peak amplitudes occur to the first input (A_1). Inputs immediately succeeding this produce a series of irregular discharges called the reorganization period (A_2). The latter period is in turn succeeded by a regular train of responses of comparatively low amplitude (A_3). Bartley (1) has given a more complete discussion of such records. The dotted line is a curve fitted by free hand to the diagram and may be compared to the brightness index changes depicted in Figure 2.

correlated to magnitude of sensory brightness. If this be so, brightness ought to decrease as the number of photic pulses in the train increases. The study to be reported tests this hypothesis by comparing the brightness of a single pulse with the brightness of pulse trains of various lengths. The variation in length of train will be such as to span the hypothetical reorganization period. Intensity and pulse-to-cycle fraction (PCF), and hence pulse length, will also be manipulated. Train separation will be fixed (held at a constant time value) since this is an important variable for brightness (4).

B. METHOD

The same episcotister and photic source set-ups used in study I (4) were used here. Also, the *O*s, psychophysical procedure for making matches, visual angles and intensity levels were as in study I.

Two large episcotister discs (four feet in diameter) were used. The periphery of each disc contained 24 equal openings equally separated. The openings of the first disc occupied $1/3$ of each cycle, and those of the second, $2/3$. Therefore when rotated in front of an appropriate source, the first disc produced trains having a PCF of $1/3$, the second, trains having a PCF of $2/3$. The relation between the sectors of the disc and slit in the lamp housing was such as to taper the pulses somewhat. Thus stimulation under the PCF $1/3$ condition was actually above zero for 37 msec. but at peak for only 29 msec. Under the PCF $2/3$ condition it was above zero for 71 msec. and at peak for 63.

Rate of rotation of the episcotister disc was fixed to give a new stimulation every 100 msec., *i.e.*, at approximately the intrinsic rhythm of the cortex which is approximately 10 phases per sec. The PCF $1/3$ condition therefore provided pulses 33 msec. in length and when more than one pulse was presented 67 msec. separated successive stimulations. Under PCF $2/3$ condition the photic pulse was 67 msec. and the interval between successive single pulses 33 msec.

Establishing a match more often than not involves a series of presentations of a given stimulus. Since the time value of null stimulation greatly affects the brightness of single pulses (4), it was necessary to carefully control the time interval separating presentations. A constant separation was achieved by blocking three successive open sectors of the large disc. The null intervals were a different constant for the two PCF conditions. PCF $1/3$ trains were all separated by 366 msec. and PCF $2/3$ trains are separated by 333 msec. Figure 2 diagrams the relationship occurring for trains in which each pulse equals 33 msec. (PCF $1/3$ condition).

C. EVALUATION OF SENSORY RESPONSE

The data are described in terms of a brightness index which is a ratio of the mean intensity of brightness matchings for the brief pulses to the mean intensity of matchings of the same target when continuously illuminated. An index of 100 represents a match when the intensities of the steady and intermittent presentations are equal; 50 a match when the steady target

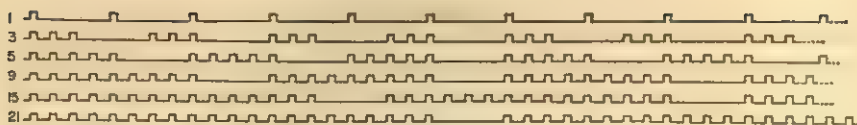


FIGURE 2

REPRESENTATIONS OF CERTAIN OF THE TEMPORAL CONDITIONS USED

The first line shows single pulses of 33 msec. followed by null periods of 366 msec. The line labelled "3" shows a train of three stimulations each 33 msec. long having separations of 66 msec. periods of no stimulation followed by a 366 msec. interval of no stimulation. The line labelled "5" shows 5 unit pulse trains separated by 366 msec. Lines "9," "15," and "21" show trains having these numbers of repetitions separated by 366 msec. null periods. Notice that, aside from line "1," the trains have 33 msec. photic periods followed by 66 msec. periods of no stimulation. Therefore, aside from the one pulse condition, all trains depicted have a pulse-to-cycle fraction of $1/3$. The pulse-to-cycle fraction $2/3$ trains used are described in the text.

had to be half as intense as the intermittent; 200, twice as intense, etc. All points above 100 therefore indicate the occurrence of brightness enhancement or the *Brücke-Bartley effect*, i.e., that the brief photic pulse is brighter than steady illumination of the same intensity. It should be pointed out in passing, however, that the index of brightness was devised simply to allow comparisons of effects induced by dissimilar conditions. The indexes are ordinal measurement and not additive in the strict sense. An index of 200 therefore does not necessarily mean that the briefly exposed target looked twice as bright as the steady but rather that the intensity of the steady target had to be doubled for the two to match. Likewise indexes of 200 do not indicate the same relative increase of brightness when the ranges of intensity matched differ.²

D. RESULTS AND DISCUSSION

Figure 3 describes the responses of the principal and student Os. Data upon which the graphs are based is given in Table 1. The differences produced by varying pulse trains are generally not large, but the values fall around the value of steady illumination (an index of 100).³

The character of the record suggests operation of a mechanism such

² In Fechnerian psychophysics, sensation changes and intensity changes have a logarithmic relation. In Stevens' scheme a power function governs the relation. In no case do the changes bear a simple linear relation to one another.

³ The brightness indexes reported here are about seven per cent higher than those of the first study under comparable conditions. This appears to have been caused by slightly higher matches for the intermittent conditions and slightly lower steady matchings. The absolute values of the matched in terms of c/ft^2 were less different than the indexes would indicate.

TABLE 1
BRIGHTNESS INDEXES ASSOCIATED WITH VARIOUS NUMBERS OF REPETITIVE IMPINGEMENTS
PCF 1/3

| Number of pulses in train | 1 | 3 | 5 | 9 | 15 | 21 |
|--|----------|-------|-------|-------|-------|-------|
| Accumulated time at average I in msec. | 33 | 99 | 165 | 297 | 495 | 693 |
| Intensity | Observer | | | | | |
| 36.3 c/ft ² | N | 127.6 | 132.6 | 109.9 | 110.3 | 109.9 |
| | BJ | 76.4 | 72.2 | 63.3 | 81.9 | 86.0 |
| 3.6 c/ft ² | N | 123.7 | 105.4 | 104.1 | 98.7 | 100.2 |
| | BJ | 78.5 | 72.4 | 63.2 | 61.1 | 66.5 |

PCF 2/3

| Number of pulses in train | 1 | 3 | 5 | 9 | 15 | 21 |
|--|----------|-------|-------|-------|-------|-------|
| Accumulated time at average I in msec. | 66 | 198 | 330 | 594 | 990 | 1386 |
| Intensity | Observer | | | | | |
| 36.3 c/ft ² | N | 196.7 | 138.0 | 137.6 | 137.1 | 122.2 |
| | BJ | 140.4 | 107.8 | 98.4 | 107.7 | 109.1 |
| 3.6 c/ft ² | N | 133.3 | 124.5 | 115.6 | 106.8 | 109.0 |
| | BJ | 90.6 | 101.0 | 83.3 | 91.3 | 84.2 |

as Bartley describes. First, decreases in brightness indexes occur as length of pulse train is increased at both levels of intensity, and for both PCFs (and hence pulse durations). Second, slopes are positively decelerated and suggestive of the curve fitted to the diagram of Figure 1. The sensory data depart from expectations only insofar as some curves of Figure 3 may decline continually over the whole range of train lengths while in the records Bartley reports reorganization was virtually completed after 11 successive shocks to the stump of the optic nerve.

The correspondence between the average value of the schematic diagram and the mean values of the sensory data is also of interest in another regard. This correspondence suggests that the matchings reflect the general charac-

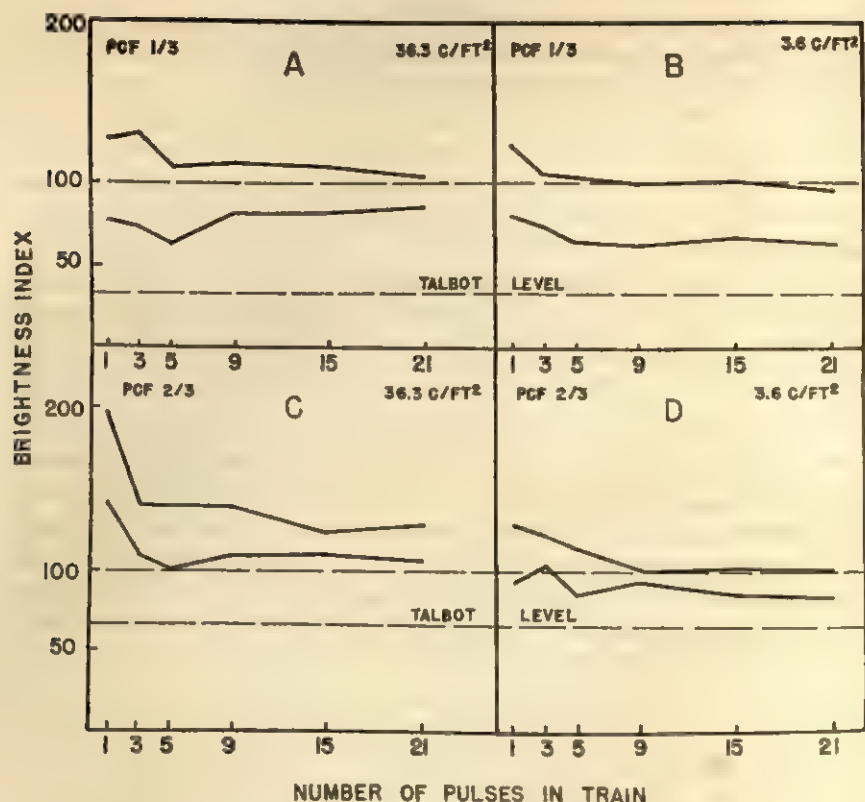


FIGURE 3
BRIGHTNESS INDEXES ASSOCIATED WITH PULSE TRAINS VARYING
IN NUMBERS OF UNITS AND PCF

A and B sections are for PCF 1/3 and C and D for PCF 2/3. A and C are for intensity levels of 36.3 c/ft.² and B and D for 3.6 c/ft.². The upper curves in each section are the mean values of the principal investigator, the lower, those for the student observers. The broken lines on the lower section of each graph represent Talbot brightness indexes and the upper broken line the value of steady illumination.

teristics of sub-fusional brightness values and are therefore like values occurring above fusion in that they refer to an average level of nervous activity during the period of stimulation. This is borne out by the fact that response variability was not systematically greater for pulse trains terminating in the reorganization period than elsewhere. Therefore matchings made during intermittent stimulation do not seem to refer to one specific stimulus among the many present (*e.g.*, the last pulse) although matching

on the basis of a single pulse such as these is clearly possible. The first points in the graph were obtained to single inputs as were many of the matching values in the previous research (4).

Finally it should be mentioned that the declines in brightness could perhaps have been made more dramatic (slope steeper) either by using longer null intervals to separate successive trains or by using higher rates of repetition for the trains. The first manipulation would have the effect of heightening the value of the single pulse but could be expected to have little effect upon the longer trains of input. The second would depress the values of the intermittent trains toward Talbot level, whereas at present the cycling is such as to produce values near enhancement (indexes greater than 100).

E. SUMMARY

The relation of brightness and number of regular repetitive stimulations was studied. At all three intensities used there was a decrease in the index of brightness as the number of elements in the stimulus train increased. This confirmed predictions made from Bartley's Alternation of Response theory and showed how mass recordings from the optic cortex may be given a sensory meaning. A feature of matching affecting the psychophysical outcome is also mentioned.

REFERENCES

1. BARTLEY, S. H. Temporal and spatial summation of extrinsic impulses with the intrinsic activity of the cortex. *J. Cell. & Comp. Physiol.*, 1936, **8**, 41-62.
2. ———. The relation between cortical response to visual stimulation and changes in the alpha rhythm. *J. Exper. Psychol.*, 1940, **27**, 624-639.
3. ———. A clarification of some of the procedures and concepts involved in dealing with the optic pathway. In Jung, R., & Kornhuber, H. (Eds.), *The Visual System: Neurophysiology and Psychophysics*. Heidelberg, Germany: Springer-Verlag, 1961.
4. NELSON, T. M., BARTLEY, S. H., & JEWELL, R. M. Effects upon brightness produced by varying the length of the null interval separating successive "single" pulses: Sensory implications of the Alternation of Response Theory, *I. J. of Psychol.*, 1963, **56**, 99-106.

Department of Psychology
Michigan State University
East Lansing, Michigan



USING ADVERBS AS MULTIPLIERS IN SEMANTIC DIFFERENTIALS*¹

University of Utah College of Medicine

ROGER B. ALLISON, JR.

A. INTRODUCTION

Cliff (2) found that adverbs tend to behave as if they were multiplying^c constants, increasing or decreasing the intensity of the words which they modify (e.g., extremely = 1.5; very = 1.3; slightly = .6). His findings were derived from testing and scaling procedures which are not unlike those used with semantic differentials and it seemed reasonable to expect that the range of response categories could be controlled through the scalar properties of adverbs. Specifically, we wanted to increase the range of the response categories of multidimensional semantic differentials and attempted to do this by changing the modifying words of the end points of the continua from "very" to "extremely." The present note reports the results of this modification when applied to the good-bad axis.

B. METHOD AND RESULTS

The data were drawn from two projects. Study I was based upon responses to a semantic differential (1) in which a series of concepts was first rated along a five point good-bad continuum and then rated along another axis. Study II consisted of responses to a different semantic differential in which each concept of the series was evaluated successively on a seven point good-bad continuum and two other continua. The two tasks had 26 concepts in common. University students ($N_I = 350$; $N_{II} = 181$) served as subjects. The numerical value of each response category for each of the concepts was calculated by Likert's method (3). The mean scale values for the categories, and corresponding standard deviations, are listed in Table 1. The reductions in variances for the tails of the two distributions were caused in part by setting arbitrary limits of ± 3.00 for the Likert values which corresponded to a percentage of zero.

* Received in the Editorial Office on March 29, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ This research was supported in part by grants from the United States Public Health Service (M-4524-C1) and Scottish Rite Foundation.

TABLE 1
SCALE VALUES OF RESPONSE CATEGORIES

| Designation of rating | Study I | | Study II | |
|-----------------------|----------|-----------|----------|-----------|
| | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> |
| Very bad | -2.52 | .58 | | |
| Extremely bad | | | -2.48 | .60 |
| Quite bad | | | -2.00 | .95 |
| Bad | -1.79 | .94 | | |
| Slightly bad | | | -1.44 | 1.02 |
| Neutral | -.68 | 1.05 | -.89 | 1.03 |
| Slightly good | | | -.35 | 1.08 |
| Quite good | | | .36 | .92 |
| Good | .47 | 1.04 | | |
| Extremely good | | | 1.43 | .67 |
| Very good | 1.66 | .83 | | |

C. DISCUSSION

It is obvious from Table 1 that the anticipated increase in range did not occur. The seven point continuum was distributed over essentially the same distance as the one based upon five categories. It seemed unlikely that the results reflected merely the variations in testing procedures between the two semantic differentials. Rather, it appeared that the subjects considered the "ends" of the continua to be the most extreme ratings, thereby ignoring the presumed differences in verbal descriptions.

These findings prompt a word of caution to investigators of semantic functioning. When semantic differentials are constructed in which the response categories are specified, the verbal labels will not necessarily constitute reference points for the subjects. It is possible that spurious differences may occur between studies and indicate only that different adverbs were used to describe the boundaries on the continua. Such dangers could be reduced if the investigators will report numerical values for the response categories.

D. SUMMARY

This study represented an attempt to enlarge the range of a response continuum for semantic differentials by using adverbs of greater intensity to describe the terminal points. The results suggested that the range is invariant with respect to such changes. It was recommended that numerical values of the response categories be reported in order to reduce potentially spurious differences in semantic differential studies which involve verbal descriptions for the response categories.

REFERENCES

1. ALLISON, R. B., JR. A two-dimensional semantic differential. *J. Consult. Psychol.*, 1963, **27**, 18-23.
2. CLIFF, N. Adverbs as multipliers. *Psychol. Rev.*, 1959, **66**, 26-44.
3. LIKERT, R. A technique for the measurement of attitudes. *Arch. Psychol.*, 1932, No. 140.

Department of Psychiatry
University of Utah
College of Medicine
Salt Lake City 15, Utah



PERFORMANCE SELECTION THROUGH RADIATION DEATH IN RHESUS MONKEYS*¹

University of South Dakota

ROGER T. DAVIS AND JAMES P. STEELE

A. INTRODUCTION

A series of investigations at Texas which were summarized by Brown and McDowell (1) indicated differences between the behavior of survivors of radiation experiments and their non-irradiated controls. These differences could have been due to direct radiation effects or could be related to selective survival which favors one kind of performance or another. The present paper describes the performance of monkeys on a variety of measures taken over a five- to seven-year span as a function of repeated irradiation with whole-body doses that were beneath the LD-50-30 and contrasts the performance of animals that survived with those that did not survive treatment.

B. METHOD

1. *Subjects*

The *Ss* were 16 rhesus monkeys procured in 1952. Each *S* was extensively trained on a variety of tasks during a period of one year, then divided at random into an experimental group of 10 *Ss* and a control group of 6 *Ss*. Each animal in the former group received 400 r whole-body radiation with X-rays on August 2, 1953, and each animal in the latter group was sham irradiated. The procedures and their effects have been described in detail (6, 13). Nine of the irradiated *Ss* survived and were irradiated for the second time with 400 r on October 16, 1954. A third period of irradiation (300 r WBR) occurred on July 18, 1956 and involved seven survivors of the experimental group. Comparisons in this paper will be made between the control group, the six animals in the experimental group that survived the longest period, and the four *Ss* in the experimental group that were the first to die.

* Received in the Editorial Office on April 2, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ This research was supported in part by research grants (M-530—M-2264) from the National Institute of Mental Health. The first author is on leave of absence at the University of Oregon, Eugene, Oregon, 1962-63.

2. Apparatus

The experiments were conducted either in the animal's living cage or in a slightly modified Wisconsin General Test Apparatus (WGTA). Radiation was administered with a 250 kvp therapy X-ray machine, and during radiation an *S* was restrained in a cylindrical tube that rotated in the X-ray beam (6).

3. Procedure

The procedures are those described earlier. Each *S* was trained on nine selected laboratory tasks. The tasks were grouped into four batteries, and after *Ss* were trained on the batteries successively, they were given a different battery on succeeding days. This procedure was employed with variations noted below, before and after each period of radiation and enabled *Es* to measure *Ss*' performance on a variety of tasks within the relatively short intervals of time during which acute radiation effects might be manifested.

Four of the tasks required the animals to make correct choices of stimulus objects for rewards of a single raisin or peanut, and six tasks demanded that *Ss* make persistent manipulative responses. The former tasks were all presented in the WGTA and included: (a) four-trial object-quality discrimination; (b) oddity principle; (c) reduced cue; and (d) delayed response. Each of these tasks was presented in a separate battery during preliminary training. The tasks involving manipulation were: (e) bent-wire detour problems with rewards of Life Savers; (f) detour problems with rewards of poker chips; (g) four simple mechanical puzzles; (h) a complex 6 unit mechanical puzzle; (i) patterned string problems; and (j) a wood block which could be chewed.

The last named task was only given before and after the first and second radiation. The remaining tasks were presented before and after the first and third period of radiation and a year after the third period of radiation. All of these tasks save the oddity problem were also presented before, during, and after a period of chronic barbiturate sedation (7). Before and after the second period of radiation three of the discrimination tasks (object-quality discrimination, delayed response, and reduced cue) and three of the manipulation tasks (patterned wire with poker chip rewards, the simple puzzles, and the wood block) were given. Reducing the number of tasks to six enabled *Es* to present the problem in two batteries and obtain measures of performance every other day in all tasks. It also afforded the opportunity to get more data on each task before and after irradiation. Although they were trained on other problems which were interspersed within the radiation

studies, Ss were always treated alike within and between different repetitions of these tasks.

C. SPECIFIC PROCEDURES AND RESULTS

1. Discrimination Tasks

The four discrimination tasks were distributed one to each battery during original training in the following order: object-quality discrimination, delayed response, oddity, and reduced cue.

a. *Object-quality discrimination.* Six 4-trial object-quality discrimination problems were given to each S daily during the 75 days of original training during September-December, 1952 and during all subsequent periods of re-training except before and after the second radiation when Ss were given four problems a day. Objects were common use or manufactured and paired at random. Stimuli used before and after the third radiation were randomly paired planometric stimuli consisting of pictures cut from magazines and pasted on 2 X 2 inch squares of Masonite.

Figure 1 shows Ss' performance on Trial 2 throughout the various periods of training. The Ss in the control group are indicated by the open circles, Ss that survived the regime of treatment by closed circles, and Ss that did not survive by closed triangles. If the performance of both segments of the experimental group were summed, the match between the experimental and control groups would be excellent. However, as the figure clearly shows, performance of Ss in the experimental group that would die was clearly superior during original training to that of the Ss that survived, $U = 0, 4, 3$, and 0 ; $p < .005, .059, .033$, and $.005$ for the sum of trials 2, 3, and 4 on days 16-30, 31-45, 46-60 and 61-75 respectively. The subgroup of experimental Ss that later died differed significantly from the control group only during days 16-30, $U = 3$, $p < .033$. Throughout the remainder of practice the experimental and control groups were significantly different only during the period one year after the third radiation wherein the performance of the Ss in the control group was significantly better than the performance of Ss in the experimental group, $U = 4.5$, $p < .017$. The performance of the animals in the group that died was not considered after the second pre-radiation period since one animal died after the first and two animals after the second period of radiation.

Several significant increments and decrements occurred between various critical periods in the experiments. There was a period of six months between the end of original training and the beginning of preradiation.

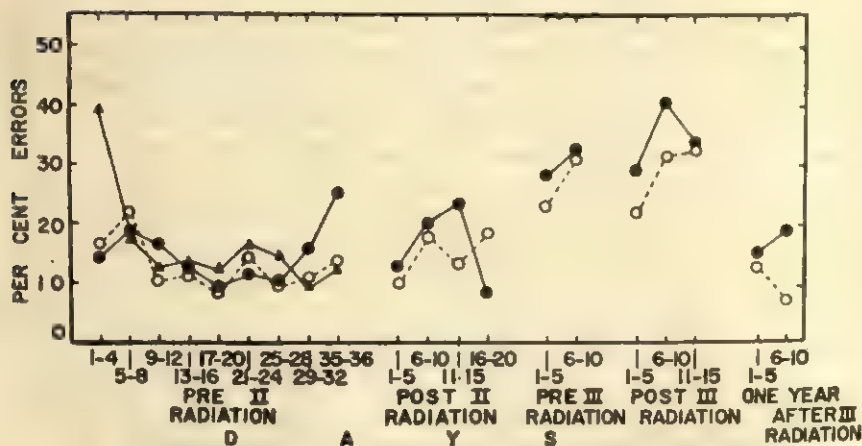
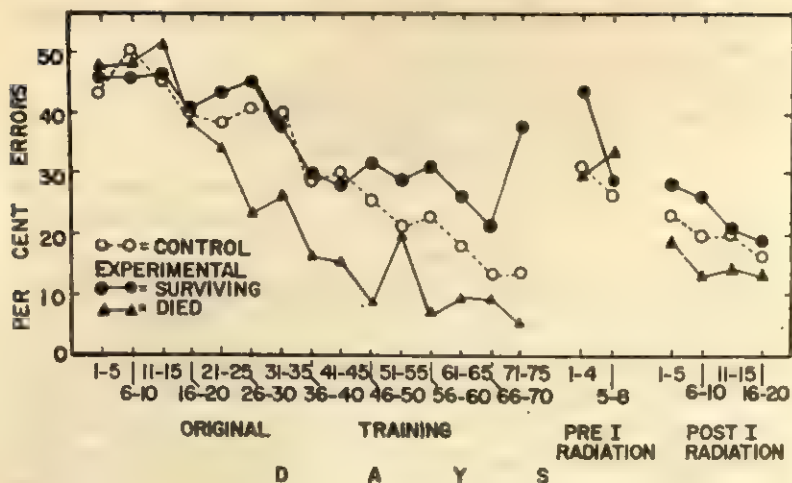


FIGURE 1
PERFORMANCE ON TRIAL 2 OF OBJECT-QUALITY DISCRIMINATION PROBLEMS
AS A FUNCTION OF TRAINING AND REPEATED TREATMENT
WITH WHOLE-BODY X-RAY RADIATION

The experimental group is divided into subgroups of Ss that survived and those that died.

training. The decrement in performance was significant and large and may be related either to forgetting or to the fact that all of the tasks were massed into four day periods preceding radiation. The decrement between the end of Postradiation II and the beginning of Preradiation III was probably

due to the fact that planometric stimuli were used before and after the third period of radiation.

b. Oddity. Subjects were given three 24-trial oddity problems a day for 30 days of original training during March and April, 1953, and during the periods before and after the first radiation and one year after the third irradiation. Before and after the third irradiation *Ss* received one 12-trial problem each day. Figure 2 clearly shows the same general relationships during preliminary training between the subjects that survived and died in the experimental group as was seen in the previous figure. The performance of *Ss* in the experimental and control groups was very well matched at the

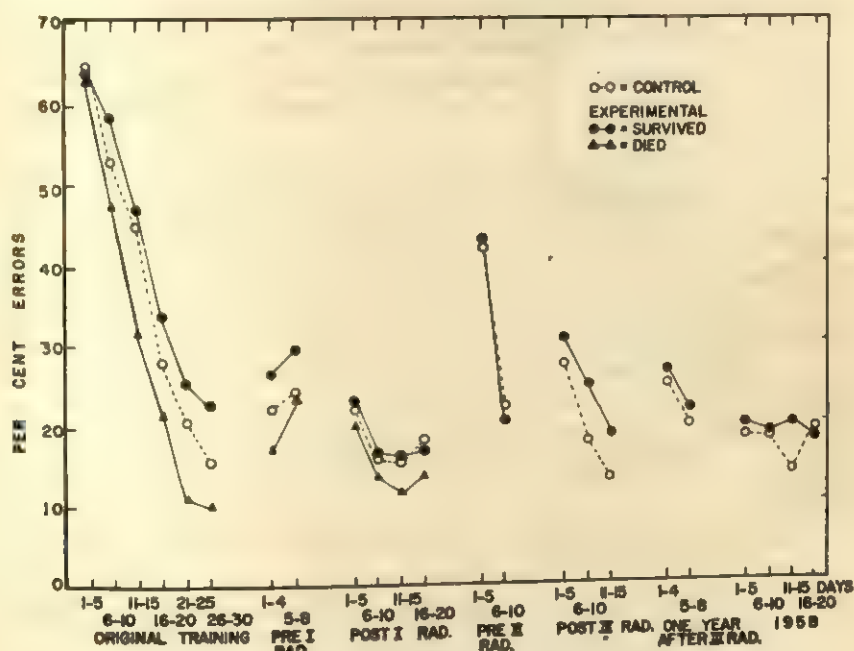


FIGURE 2
PERFORMANCE ON ODDITY PROBLEMS AS A FUNCTION OF TRAINING
AND REPEATED RADIATION

end of preliminary training, but during the last two periods of preliminary training, days 21-25 and 26-30, the *Ss* in the experimental group that died after radiation made significantly fewer errors than *Ss* in the experimental group that survived, $U = 0$, $p < .005$, $U = 2$, $p < .019$ respectively.

The performance of the *Ss* remained stable throughout the rest of training

except at the beginning of the period preceding the third radiation when all *Ss* made more errors than at the end of the period following the first radiation, $p < .01$, using the sign test. By the second five-day period preceding the third radiation *Ss* returned to a level comparable and not significantly different from the last five days following the first radiation.

c. Reduced cue. The procedures of reduced cue problems were a compromise between the procedures of delayed response and discrimination (2). Each problem contained two trials, a cue trial and a reduced cue trial. Identical objects were used on both trials to cover the two foodwells in the stimulus tray of the WGTA. During the cue trial a third object was placed on top of one of the identical objects to indicate that it was rewarded. The cue was removed during the second trial and *S* was rewarded for displacing the object that contained the cue on the first trial. The procedures for manipulating the WGTA were identical to those used in presenting discrimination problems.

Each *S* received 24 problems a day except during the 36 days preceding and the 20 days following the second radiation. The exceptions consisted of giving 40 trials a day, divided into four conditions. These conditions were presented in a 2×2 factorial design in which half of the trials involved a familiar pair of objects and half involved an unfamiliar pair. Half of the trials using familiar and half using unfamiliar objects were presented in the manner described above. The remaining half required *S* to respond on the reduced cue trial to the object and position that was not rewarded on the cue trial. The former procedure had no appreciable effect, and this data was pooled. The latter procedure was insoluble, and this data was not used.

During original training the number of incentives was varied on the cue and reduced cue trials, but this operation did not change performance significantly and was not employed thereafter.

The performance of *Ss* shown in Figure 3 also resembled that shown in previous figures. The segment of the experimental group that died differed significantly from the segment that survived, $U = 0$, $p < .005$ during the days 5-8 in the original training. During days 9-12 of original training and days 5-8 of preradiation training, the difference in performance of these two subgroups approached significance, $U = 4$, $p < .057$ in each case. At no point in practice did the experimental and control groups differ significantly, and the consistently better performance of *Ss* in the control group after the second radiation is due to two animals that made an excessive number of errors.

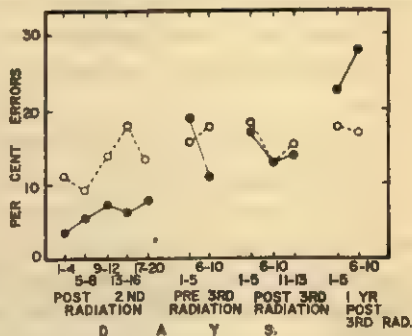
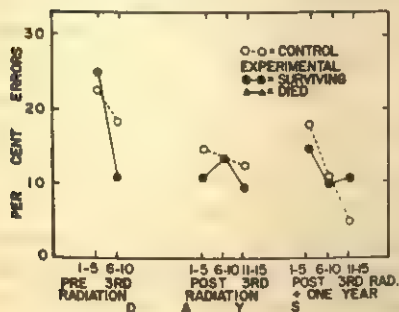
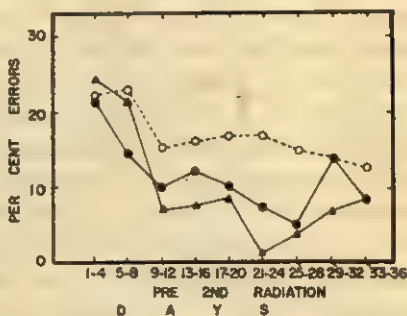
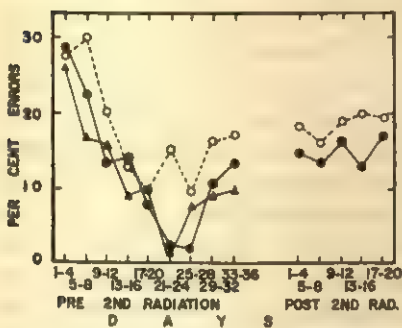
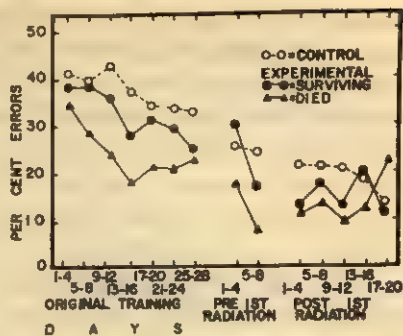
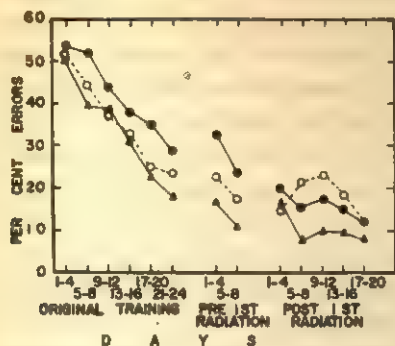


FIGURE 3 (left column)

PERFORMANCE ON REDUCED CUE PROBLEMS AS A FUNCTION OF TRAINING AND REPEATED RADIATION

FIGURE 4 (right column)

PERFORMANCE ON FIVE-SECOND DELAYED RESPONSE PROBLEMS AS A FUNCTION OF TRAINING AND REPEATED RADIATION

d. Delayed response. Each *S* was given 16 delayed response trials a day during the 28 days of preliminary training and during the period one year after the third radiation. Before and after the second radiation each *S* received 10 trials a day, and before and after the third radiation 20 trials were given each day. Half of the trials given each day in preliminary training employed five-second delays, and the other half employed 20-second delays. Thereafter, only five-second delays were used.

Figure 4 shows the performance of the *Ss* in the control group and the two divisions of the experimental group throughout the series of experiments. Matching of the experimental and control groups was less perfect for the delayed response problem than for the other three discrimination problems; however, none of the differences between the groups during preliminary training were significant, although several of the points differed at the 10 per cent level. The difference between the performance of the animals that died later in the experiment and those that survived was due to the fact that two of the *Ss* in the former subgroup performed much more efficiently than any of the other *Ss* and the other two were similar in their performance to the remainder of the *Ss*. During the postradiation period following the second radiation the performance of the *Ss* in the experimental group that survived and the control group was significantly different, $U = 7$, $p < .047$. After this period the performance of the animals that had been irradiated deteriorated markedly and became inferior to that of *Ss* in the control group. This deterioration was significant, $U = 6$, $p < .05$. During the same period the performance of the animals in the control group remained relatively stable.

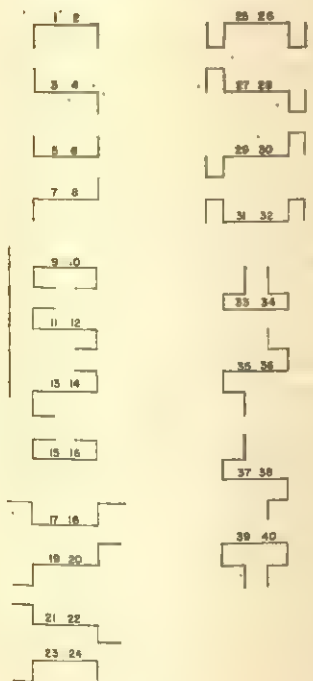
2. Manipulation Tasks

During original training *Ss* received three manipulation tasks in the first battery: bent-wire problems with Life Saver rewards, patterned string tests, and the 6-unit mechanical puzzle. In the second battery of tasks they received the bent-wire problem with poker chip rewards, and simple puzzles. The third battery did not contain any of the manipulation tasks, and the fourth battery contained the wood block test.

a. Bent-wire detours with Life Saver rewards. The methods and results of the studies with bent-wire detour problems were reported in detail (6, 8). The stimulus patterns are shown in the left hand side of Figure 5. They were constructed of 0.187-inch welding rod and consisted of a center piece six inches long and one or more two-inch pieces bent 90° either forward or backward at either end of the center segment. Patterns differed in difficulty

according to the direction *S* pushed the lure and the number of segments in the pattern. Each *S* received seven patterns a day during 74 days of preliminary training and two on the 75th day. This represented 13 replications of 40 patterns. Another replication was presented immediately before the first radiation and three replications immediately after radiation. Twelve

BENT WIRE PATTERNS USED WITH LIFE-SAVER
REWARDS



BENT WIRE PATTERNS USED WITH POKER CHIP REWARD



FIGURE 5
BENT-WIRE PROBLEMS USED WITH REWARDS OF "LIFE SAVERS"
AND PUNCHED PAPER POKER CHIPS

representative patterns, Nos. 1-2, 5-6, 17-18, 23-24, 25-26, 31-32, were presented on 10 days before and 15 days after the third radiation. One year after the third radiation all 40 patterns were presented one time each in the manner used in the period of original training.

Figure 6 shows the median speed to remove the lure for *Ss* in each of the three groups on five different levels of complexity of pattern. Heretofore the data was presented in terms of the number of failures (6), but after

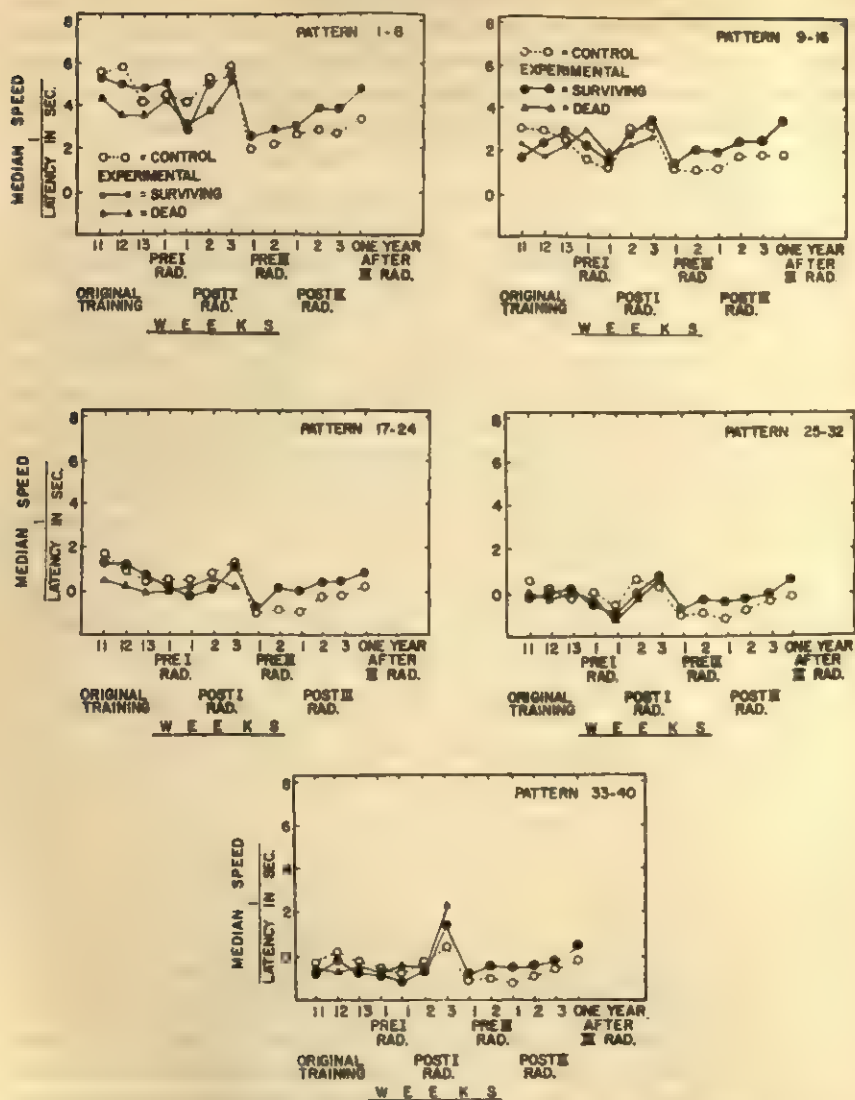


FIGURE 6

MEDIAN SPEED REQUIRED TO REMOVE "LIFE-SAVER" REWARDS FROM BENT-WIRE DETOUR PROBLEMS AS A FUNCTION OF TRAINING AND REPEATED RADIATION

the 11th week of training the number of failures of *Ss* to remove the candy from the pattern approached zero, and the use of latency or its reciprocal, speed, was more appropriate. During the last six replications in original training the median speeds of *Ss* in the two treated groups, those that survived and those that did not, were less than speeds of *Ss* in the control group. However, in 30 comparisons—five levels of complexity in six replications—only three differences between the performance of *Ss* in the surviving experimental group and the control group were significant. On patterns 9-16 during replication 11, the experimental group that survived plus those that died were significantly different from *Ss* in the control group.

No points are plotted for patterns 9-16 and 33-40 between the end of the first postradiation period and one year after the third radiation period because during the intervening period an abbreviated group of tests was employed. It is very clear that retention within this period was poorer on all patterns for *Ss* in the control group than *Ss* in the experimental group that survived, and this is reflected in significant differences between the groups during the period that preceded the third radiation, days 6-10 on patterns 17-18, 23-24, and 25-26; $U = 0$, $p < .037$, $U = 0$, $p < .001$, $U = 6.5$, $p < .021$ respectively, $n_1 = 6$, $n_2 = 7$. The performance of *Ss* in the experimental group that died did not seem to affect this result. Two of these *Ss* performed very well and two very poorly on these problems, and their death only served to make the data more homogeneous.

b. Six-unit puzzle. This puzzle was similar to one used in Harlow's laboratory (9) and consisted of two metal pins, a doorstop, a bolt, a hasp, and a hinge, which can be displaced in the order given. The puzzle was attached to the bars in the center of the door of the living cage during a 15-minute period each day for 60 days during original training.

The *Ss* learned very quickly to displace the elements, and as reported earlier (6) the performance of the groups reached asymptotes between 4.0 and 4.5 elements per puzzle presentation. Individual *Ss* were very consistent in performance, and some *Ss* reached asymptotes which represented displacement of 5.5 to 6.0 elements while others displaced as few as 2.5 elements per presentation. Significant changes as a function of practice occurred only during preliminary training and no significant changes were attributable at any segment of the task to forgetting or to the group to which *S* was assigned or ordained.

c. Patterned string tests. Six different patterned string tests were presented to each animal in a random order daily for 75 days during original training and before and after the first and third periods of irradiation and

the testing one year after the last period of irradiation. Strings were constructed of different lengths of plumbers chain with a harness snap at one end and a small piece of wire at the other end. A raisin or grape impaled on the wire served as a reward and several strings were presented together in patterns with one string reinforced. The monkey responded by pulling on the end of the string that was snapped to one of a series of screw eyes that were equally spaced along the edge of the board on which the patterns were presented. The patterns, which have been described in detail (5) were: cross strings, box pattern, angle cross pattern, four crossing strings, double cross, and pseudo cross.

Figure 7 shows the performance of Ss in the control group and the two parts of the experimental group on patterned string tests as a function of training and treatment. When the data was analyzed at the end of the original training, the differences between the experimental and control groups were not significant. However, elimination of the animals that later died made the difference between the experimental and control groups significant during days 26-50, $U = 5$, $p < .02$ and days 51-75, $U = 6$, $p < .03$ respectively. After preliminary training the performance of Ss deteriorated.

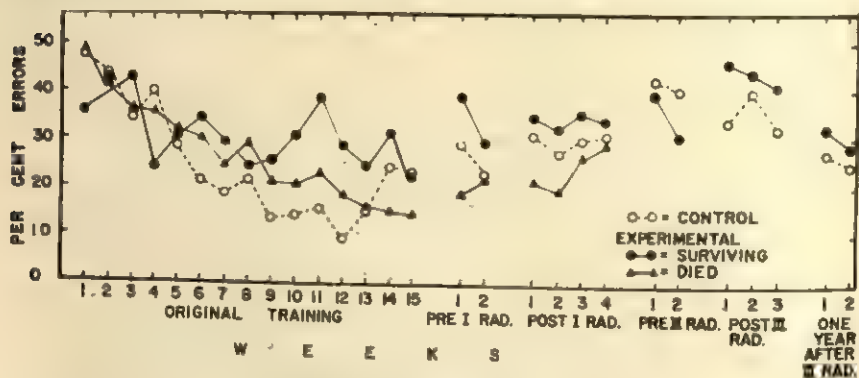


FIGURE 7
PERFORMANCE ON SIX PATTERN STRING TESTS AS A FUNCTION OF TRAINING
AND REPEATED RADIATION

The decrement in performance of Ss between the last third of original training and the training before and after the third radiation was significant using the sign test, $p < .05$ and $< .01$ respectively. In the training before the third radiation the animals in the experimental group performed significantly better than Ss in the control group, $U = 3.5$, $p < .01$. Between the pre- and postradiation training of the third radiation and performance

of Ss in the control group improved and Ss in the experimental group deteriorated. An analysis of variance of the data before and after the period of the third radiation indicated that the interaction between *treatment group* and *treatment* was significant, $F = 14.06$, $p < .01$ using the interaction *within groups* \times *treatment* as the error term.

d. *Simple puzzles.* Four manipulation devices were presented for two minutes each day to Ss in their living cages and in the WGTA. Since manipulation was much more frequent in the living cage than in the WGTA, only the data collected in the former site will be considered.

The first device consisted of nine 2.5-inch screw eyes arranged three inches apart along the length of the board and two inches apart along the width of the board. Device 2 was composed of nine pairs of two-inch hooks and eyes arranged as the first device, and the third device included nine hasps and hasp staples mounted in the same spatial arrangements as the first and second devices. Device 4 was a composite of the elements in the first three devices and consisted of three successive sets of units. Each set included a hasp restrained by a hook which in turn was restrained by a screw eye. Each S received four days of original training with each of the puzzles in the living cage, four days during the training that preceded the first period of irradiation, ten days during the 60 calendar days that followed radiation, and ten days during a period one year after the third radiation. During the latter period only the three puzzles with independent elements were presented.

There were no significant changes in performance on simple puzzles other than those reported earlier, a transitory drop in the manipulation of the second puzzle between the periods before and after the first radiation by Ss in the experimental group (6) and complex changes during sedation (7).

e. *Bent-wire poker chip task.* Each of the Ss received six bent wire problems each day during 32 days of original training and on the days previously designated before and after each of the three periods of irradiation and one year after the third radiation. Every two days all of 12 patterns which were described previously (6) were presented to Ss in their living cages. Four of these patterns were identical to those used with rewards of Life Savers, and eight were more complex. (See Figure 5.) Reinforcement was provided by punched cardboard poker chips placed on the bent wire pattern against jaws of a vise grip which was held in position within the living cage by means of a large C-clamp. Time taken to remove the poker chip was recorded, and the trial was terminated as a *failure* if the poker chip was not

removed within 45 seconds. Differentiation was made between *failed without trying* and *failed with trying* and between manual and oral manipulation.

The only finding that has not heretofore been reported is that the animals in the experimental group tended to succeed in removing the poker chip significantly more frequently by chewing than *Ss* in the control group, $U = 0$, $n_1 = 6$, $p < .001$, during the original training period.

The significant difference between the experimental and control group that was reported earlier (6) occurred during the period of the first radiation and was restricted to the period of the sixth replication of patterns after radiation, i.e., calendar days 12 to 16 after radiation. The two parts of the experimental group responded in a similar manner.

f. Wood-block chewing task. Each *S* was given a block upon which to chew in its living cage for 15 minutes each day while the cagemate was being tested in the WGTA on the reduced-cue problem. A block consisted of two $3 \times 3 \times 1$ inch pieces of pine board bolted together. Blocks were weighed nightly and a particular block was always given to the same *S*. New blocks were given each time a new phase of training commenced and used throughout the particular phase. Blocks were given during original training and before and after the first and second radiation.

As previously reported (6) *Ss* in the control group showed a significant spurt in chewing after the first radiation followed after the 16th post-radiation day by *Ss* in the experimental group. A comparison between *Ss* in the experimental group that were to survive and die was not significant, nor were differences between the experimental and control group preceding or following the second radiation.

3. Task Reliability and Interrelatedness

Considering that the various tasks could be grouped according to the similarity of effects, *Es* determined the rank order correlations between performance on the ten tasks during the period that immediately preceded the first radiation. This period was selected because the tasks had all been learned previously, all *Ss* were alive, none had received treatment, and all tasks were presented to *Ss* within four-day periods of time. The findings are similar to those of Harlow (10) who reported high intratask and low intertask correlations. In the present study, tasks that employed food rewards, with the exception of oddity, tended to be positively intercorrelated—see bent-wire Life Saver, pattern strings, object-quality, reduced cue, and delayed response. However, none of the correlations were very high

and only a few were significant. The correlations between tasks that did not involve food correlated negatively or only slightly with tasks that did. However, performance on simple puzzles correlated significantly with reduced cue problems.

D. DISCUSSION

The results of these investigations indicated large and significant differences between the performance of monkeys that have received and survived three widely spaced doses of radiation with X-rays and monkeys that have not been irradiated. Furthermore, differences were found in the present study between the performance of the survivors of the three periods of radiation and their controls on tasks that did not differentiate between irradiated and nonirradiated Ss in earlier studies (6, 11, 14). However, transient changes that followed the first radiation (6) were not evident after the second and third radiations.

There were two types of change in behavior in the present study that were attributable to the effects of radiation. One involved selective survival of Ss in the experimental group and the other involved progressive changes in the behavior of Ss that had been irradiated.

1. *Selective Survival*

Performance of Ss in the experimental and control groups was adequately matched at the end of the preliminary training on oddity and object-quality discrimination problems. However, the performance of Ss in the experimental group *that survived*, differed significantly from the performance of Ss in the control group during preliminary training, and differences between the groups which persisted throughout three periods of radiation were due to selection through death rather than progressive debilitation. Significant selection also occurred in the patterned string task and selection that was not significant was suggested by performance in several other tasks.

2. *Progressive Radiation Effects*

Four tasks illustrated radiation effects: bent-wire detour problems with Life Saver rewards, delayed response, reduced cue, and patterned string tests. For the first three of these tasks death tended to increase the similarity of the experimental Ss that survived and the control Ss, but death tended to separate the performance of the groups on patterned string problems.

Performance of monkeys in the subgroup of the experimental group that survived is somewhat, although not significantly, superior to the performance

of Ss in the control group during preliminary training on bent-wire detour problems with rewards of Life Savers, but ultimately during the period one year after the third radiation Ss in the control group were inferior in performance to Ss in the experimental group that survived. These changes were especially noteworthy because this task clearly distinguished the performance of Ss of different groupings of species, age, and ability (8, 15, 17).

The performance of Ss in the group of survivors was superior to that of Ss in the control group throughout most of the training on delayed response problems. However, the performance of Ss in the former group deteriorated after the second period of irradiation and by one year after the third radiation the performance of Ss in the experimental group was significantly inferior to the performance of Ss in the control group. These results only partially resemble the findings of Brown and McDowell (1) who reported that chronic survivors of experiments with large doses of whole-body radiation with a mixed source perform less well on delayed response problems than Ss that have received a somewhat lower dose. However, contrary to the trends of the present experiment, they found that the Ss in their control group were the least proficient of the three groups that they studied.

The changes in performance of both reduced cue and patterned string problems were complex but nevertheless clearly related to radiation effects. Brown and McDowell (1) reported a facilitative effect of radiation on reduced cue discrimination problems, which agrees with the overall trend of the present experiment.

The next obvious question to consider is whether these various findings contribute to the understanding of the radiation syndrome in monkeys. Brown and McDowell (1) parsimoniously concluded that radiation produced decreased distractibility and increased concentration of attention. They further concluded that irradiated Ss did better than nonirradiated Ss on tasks that required Ss to attend to objects. Two of the tasks that were used in the present experiments, object-quality discrimination and oddity, require recognition and discrimination of the qualities of the objects. Efficient performers on these same tasks died selectively in the present experiments. These Ss were typically hypoactive, attentive, and apparently less viable than the other Ss. The findings of the present experiment, although in the same direction as those of the workers at Texas, indicate that the decrement in performance was not progressive, and this contraindicated the notion that the change was related to tissue damage.

Tasks in the present experiment that required special attention to food

by Ss were delayed response problems and bent-wire detour problems with Life Saver rewards. The Ss that survived radiation did better on bent-wire problems with Life Saver rewards than the control Ss during the final period of training in spite of differences in the opposite direction during preliminary training. This corroborated the generalization made by Brown and McDowell, but their conclusions were refuted by the decrement in performance of animals in the experimental group on delayed response problems. This lack of agreement between the findings of the present studies and the conclusions made by Brown and McDowell together with the demonstration that apparent decrements in performance could be due to selective rather than progressive damage and the earlier suggestion that the distractibility effect may be motivational (4) indicates that further theoretical treatment of the syndrome of whole-body radiation is necessary.

E. SUMMARY

The performance of two groups of monkeys was studied over a seven-year span on selected laboratory tasks. One group received three doses of whole-body radiation with X-rays in near lethal doses. Four of these Ss died within the span of the experiments, and six survived. The performance of the surviving and nonsurviving Ss was contrasted with the performance of the six Ss that comprised the nonirradiated control group.

1. The immediate effects of the first period of irradiation involved tasks that had a high component of manipulation (6). These changes were transitory and did not occur in subsequent postradiation periods.

2. When Ss that died were removed from consideration, the originally good matches between the performance of treated and untreated groups disappeared on object-quality discrimination and oddity tasks, and the performance of the Ss that survived treatment was significantly inferior to the performance of the nonirradiated Ss. This suggested that death selected irradiated Ss that were efficient on these tasks.

There were chronic changes in performance on four tasks. The Ss in the control group had a decrement in performance on bent-wire detour problems with Life Saver rewards; Ss in the experimental group showed a decrement in performance on delayed response problems; there was a significant interaction between pre- and postradiation performance of Ss in the experimental and control groups on patterned string tests that were given before and after the third radiation; and radiation apparently facilitated performance of reduced cue problems.

REFERENCES

1. BROWN, W. L., & McDOWELL, A. A. Some effects of radiation on psychological processes in rhesus monkeys. In Haley, T. J. and Snider, R. S. (Eds.), *Response of the Nervous System to Ionizing Radiation*. New York: Academic Press, 1962.
2. DAVIS, R. T. The performance of monkeys on discrimination problems with reduced stimulus cues. *Proc. S. Dak. Acad. Sci.*, 1954, **33**, 116-122.
3. DAVIS, R. T., ELAM, C. B., & McDOWELL, A. A. Latent effects of chronic whole-body irradiation of monkeys with mixed source radiation. U.S. Air Force School of Aviation Med. Report No. 57-59, 1958, 1-31.
4. DAVIS, R. T., & LOVELACE, W. E. Variable rewards and peripheral cues in discriminations by irradiated and nonirradiated monkeys. *J. Genet. Psychol.*, in press.
5. DAVIS, R. T., & McDOWELL, A. A. Performance of monkeys on randomly presented string problems. *Proc. S. Dak. Acad. Sci.*, 1953, **32**, 147-152.
6. DAVIS, R. T., McDOWELL, A. A., DETER, C. W., & STEELE, J. P. Performance of rhesus monkeys on selected laboratory tasks presented before and after a large single dose of whole-body X-irradiation. *J. Comp. & Physiol. Psychol.*, 1956, **49**, 20-26.
7. DAVIS, R. T., McDOWELL, A. A., GRODSKY, M. A., & STEELE, J. P. The performance of X-ray irradiated and nonirradiated rhesus monkeys before, during and following chronic barbiturate sedation. *J. Genet. Psychol.*, 1958, **93**, 37-51.
8. DAVIS, R. T., McDOWELL, A. A., & NISSEN, H. W. Solution of bent-wire detour problems by monkeys and chimpanzees. *J. Comp. & Physiol. Psychol.*, 1957, **91**, 233-238.
9. DAVIS, R. T., SETTLAGE, P. H., & HARLOW, H. G. Performance of normal and brain operated monkeys on mechanical puzzles with and without food incentive. *J. Genet. Psychol.*, 1950, **77**, 305-311.
10. HARLOW, H. F. Behavioral contributions to interdisciplinary research. In Harlow, H. F., & Woolsey, C. N. (Eds.), *Biological and Biochemical Basis of Behavior*. Madison, Wisconsin: U. Wisconsin Press, 1958.
11. HARLOW, H. F., & MOON, L. E. The effects of repeated doses of total-body X-radiation on motivation and learning in rhesus monkeys. *J. Comp. & Physiol. Psychol.*, 1956, **49**, 60-65.
12. HOLLIS, J. H. Solution of bent-wire problems by severely retarded children. *Amer. J. Ment. Defic.*, 1962, **67**, 463-472.
13. McDOWELL, A. A., DAVIS, R. T., & STEELE, J. P. Application of systematic direct observational methods to analysis of the radiation syndrome in monkeys. *Percept. & Motor Skills*, 1956, **6**, 117-130.
14. RIOPELLE, A. J., GRODSKY, M. A., & ADES, H. W. Learned performance of monkeys after single and repeated X-irradiations. *J. Comp. & Physiol. Psychol.*, 1956, **49**, 521-524.
15. WHITECRAFT, R. A., COBB, H. V., & DAVIS, R. T. Supplementary report: Solution of bent-wire detour problems by preschool children. *Psychol. Rep.*, 1959, **5**, 609-611.

Department of Psychology
 University of Oregon
 Eugene, Oregon

THE CORRELATION OF THE PORTEUS MAZE
AND THE GESTALT CONTINUATION
AS PERSONNEL SELECTION TESTS
OF PERIPHERAL PEOPLES*

University of Hawaii

A. JAMES GREGOR AND D. ANGUS MCPHERSON

A. INTRODUCTION

The unprecedented industrialization which has characterized the Twentieth Century has mobilized to its service peripheral peoples in the furthest reaches of the globe. Such has been particularly the case in Africa and more recently in Central Australia. South African psychologists have been particularly active in devising, administering, validating and standardizing personnel selection tests to be employed in the aptitude assessment of nonliterate Ss for mechanical, nonmechanical and supervisory tasks (1, 2, 3, 4, 5, 6, 7). In the United States the Porteus Maze has been effectively employed in similar assessments of potential, particularly among those academically impaired (in the case of Jensen's study, primarily of the Negro subculture [9, 10]).

Field work among the nonliterate Aborigines of Central Australia¹ permitted the administration of both the Porteus Maze and the Gestalt Continuation pencil and paper selection tests to the same subject population in the effort to search out an implied affinity between the two measures. Both tests purport to test a factor of general intelligence neglected by standard academically orientated tests but serviceable in personnel selection for industrial tasks and as a prognostic index of general adaptability (8, 11, 13). On the other hand, there is a substantial and fundamental difference in the character of the tests and the nature of their demands upon the S. The Maze Test has proven itself one of the most durable psychometric devices available to psychologists, having unique diagnostic and prognostic merit (13). The Gestalt Continuation Test (8, 14) is an extremely simple pencil and

* Received in the Editorial Office on April 8, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ The tests were administered by the senior author in the course of field work under the direction of Professor Emeritus S. D. Porteus with the assistance of a generous grant from the Human Genetics Fund. The authors are solely responsible for the exposition, analysis, and conclusions of this report.

paper test, validated with a correlation of $r = .48$ (significant at the one per cent level) with the South African General Adaptability Battery (8).

B. METHOD


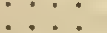
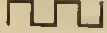


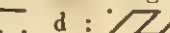
1. Subjects

Ss were forty-one males, 18 to 45 years old, of the Wailbri and Pintubi aboriginal tribes of Central Australia. All were illiterate and spoke little English. All, with the exception of one, readily submitted to testing.

2. Materials

Three batteries (Vineland Revision, Extension and Supplement) of the Porteus Maze Test were administered followed by the administration of the Gestalt Continuation Test (Form A).

3. Procedure

The Porteus Maze Test, in three batteries, was administered individually in total test time ranging from sixty to one hundred and eighty minutes. A rest period was permitted between batteries. Standard testing procedures for Maze administration were followed throughout. Directions were conveyed by mime and example and Ss executed preliminary practice mazes in accordance with established testing technique. Upon completion of the Maze Test, the Gestalt Continuation Test was administered in a test time averaging approximately five minutes. A select pattern was shown the S () and the experimenter, using a soft pencil, traced the printed part of the pattern continuing the action, without putting the pencil to paper, for one further unit into the two straight lines formed by regularly spaced dots (). The Ss observed how the pattern was traced and were directed, by mime, to complete the pattern to the right. The test patterns to be executed were in four variations of differing complexity (a :  b :  c :  d : ). One modification in the administration of the Gestalt was found necessary. When the tester traced the printed pattern over the page without touching pencil to paper it was found that some of the first Ss repeated the performance, i.e., not touching pencil to paper. As a consequence, the administrator carried the test pattern one unit into the right actually fashioning one further section. "Insight" was immediately evident.

The tests were administered either on the grounds of the Native Settlement at Yuendumu and Papunya on a table provided by the resident staff, or

in situ, in the native encampments themselves, on a piece of plywood which the senior author transported there for that purpose.

C. RESULTS

A correlation of $r = .73$ was found to obtain between the Porteus Maze and the Gestalt Continuation Test in a subject population of 40 male Aborigines. The scores of one S were not used in this r as he only reluctantly submitted to testing. His disposition was characterized in the test records as "sullen," and it proved impossible to establish rapport. His row score (Porteus Maze) deviation was significant beyond the .05 level and his column score (Gestalt Continuation) deviation just short of significance at the .05 level. With the inclusion of his scores, the correlation obtained is reduced to $r = .63$.

D. DISCUSSION

Considering the extreme simplicity of the Gestalt Continuation Test the fact that a correlation of .73 obtains with respect to the well established Porteus Maze is of considerable interest. One draws conclusions from such a finding, given the size and character of the test population and the obvious differences between the test devices, only at considerable hazard. Since, however, both tests claim some success in trainee selection (well-validated in the case of the Porteus Maze), and our research with a small sample of primitive subjects indicates a correlation of a reasonably high order, there seems to be some reason to believe that they are testing, at least in part, something of the same order. That this is, in fact, the case could only be determined by further empirical studies, ideally validating the tests for such Ss independently against some objective index. Whether such a correlation would obtain with more sophisticated groups seems doubtful.

Results obtained on the G.C.T. from this sample of Aborigines seem to indicate that the test is impaired by inadequacies in scope and inappropriateness in the configurations of patterns selected, and scoring method now being used (14)—at least with respect to such a group as that herein considered. The marginal distribution (Figure 1) shows a strong tendency toward bimodality which, in part, is a consequence of four factors: (a) the failure of scores on individual items to sort out over the possible range of .00 to 2.00; (b) a high negative correlation² between certain items (a & d and b & c); (c) a bonus system of scoring which aggravates the clustering of scores by selectively introducing discontinuities into the score continuum (hatched

² $\phi_{ad} = -.49$, $\phi_{max.} = -1.00$; $\phi_{bc} = -.47$; $\phi_{max.} = -.89$. $P > .01$.

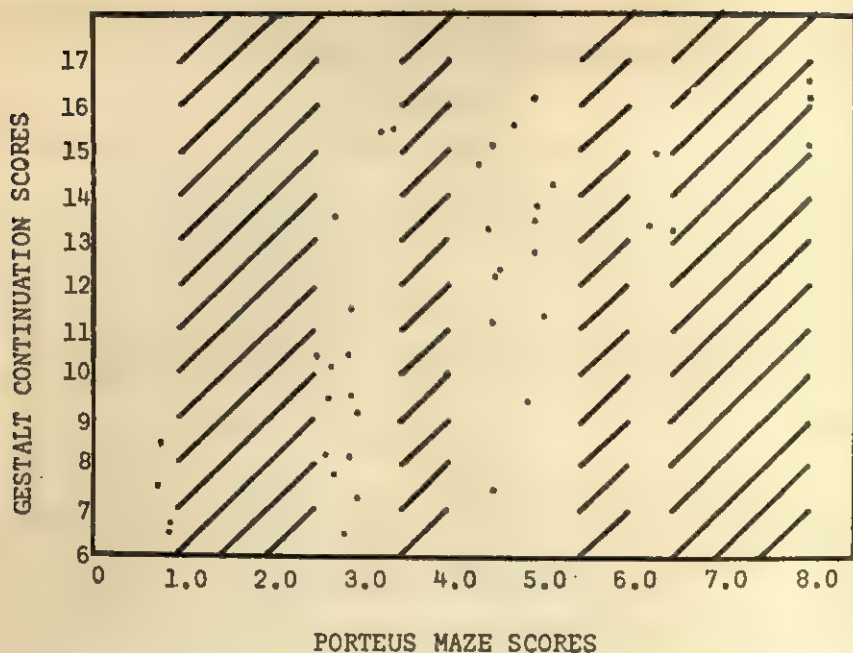




FIGURE 1

SCATTER DIAGRAM OF PORTEUS MAZE AND GESTALT CONTINUATION SCORES ILLUSTRATING DISCONTINUITIES AND BIMODALITY OF GESTALT DISTRIBUTION

areas of Figure 1); (d) systematic errors in pattern continuations stemming from the similarities of patterns presented on the test (a & d and b & c) and, perhaps, abetted by the "setting" of early patterns due to the almost total unfamiliarity of the Ss with the use of a pencil, in conjunction with a larger number of completions per pattern than are necessary to establish the S's ability (i.e., having executed , and being confronted with , there is a strong tendency for the S to revert to the first pattern). Should these sources of error represent extraneous "contaminations" they will, so long as their effects remain undetermined and/or uncompensated for, serve only to reduce unnecessarily the sensitivity of the G.C.T.

There are a host of other problems attending the use of such tests with peripheral peoples, both of a theoretical and practical nature, which cannot be treated within the compass of this preliminary report. As long as these problems are unresolved, it is an open question as to whether or not the specific results obtained from this aboriginal group may be viewed as pre-

dictive of those apt to be obtained with others. The evidence does indicate, however, that pencil and paper tests may conceivably be used in the personnel selection requisite to the needs of basic industrial training which will, for at least the immediate future, constitute the principal task confronting those charged with the uplifting of peripheral peoples.

E. SUMMARY

Forty-one adult male Wailbri and Pintubi Australian Aborigines were tested using the Porteus Maze Test (Revision, Extension and Supplement) and the Gestalt Continuation Test (Form A), the tests being employed as devices for personnel selection of nonliterates. It was impossible to establish rapport with one S, the correlation obtained using the results from the remaining 40 being $r = .73$. The small sample and the singularity of the population make interpretation hazardous, but the extreme simplicity of the G.C.T. and its reasonably high correlation with the well established Porteus Maze make the results an inducement for further studies of the pair. The results obtained from this sample seem to indicate that the G.C.T. is unnecessarily impaired by inadequacies in scope and inappropriateness both in the pattern configurations employed and in the scoring system now being used.

REFERENCES

1. BIESHEUVEL, S. Psychological tests and their application to non-European peoples. In *The Yearbook of Education*. London: Evans, 1949.
2. ———. The occupation abilities of Africans. *Optima*, 1952, 2, 18-22.
3. ———. Personnel selection tests for Africans. *S. Afr. J. Sci.*, 1952, 49, 3-12.
4. ———. The study of African ability. *Afr. Stud.*, 1952, 11(2), 45-48 & 11(3), 105-117.
5. ———. The measurement of occupational aptitudes in a multiracial society. *Occup. Psychol.*, 1954, 28, 1-8.
6. ———. Objectives and methods of African psychological research. *J. Soc. Psychol.*, 1958, 47, 161-168.
7. ———. Research into the occupational fitness of Africans in the Union of South Africa. *Bull. Inter-Afr. Labour Inst.*, 1957, 4(5), 1-15.
8. HECTOR, H. Results from a simple Gestalt Continuation Test applied to illiterate black mineworkers. *J. Nat. Inst. Person. Res.*, 1960, 8, 145-147.
9. JENSEN, M. B. Basic airmen who fail to meet minimum mental requirements in the air force. In *Lackland Air Force Base Report*, November, 1952.
10. ———. The "low level" airman in retesting and basic training: A sociopsychological study. *J. Soc. Psychol.*, 1961, 55, 177-190.
11. PORTEUS, S. D. A survey of recent results obtained with the Porteus Maze Test. *Brit. J. Educ. Psychol.*, 1952, 22(3), 180-188.
12. ———. What do the Maze Tests measure? *Aust. J. Psychol.*, 1958, 10(3), 235-256.

13. ———. *The Maze Test and Clinical Psychology*. Palo Alto: Pacific Books, 1959.
14. TEKANE, I. A new and objective scoring method for the Gestalt Continuation Test. *J. Nat. Inst. Person. Res.*, 1960, 8, 148-150.

Department of Philosophy
University of Hawaii
Honolulu 14, Hawaii

TYPES OF EMOTIONS OR DIMENSIONS OF EMOTION?
A COMPARISON OF TYPAL ANALYSIS WITH
FACTOR ANALYSIS*†

Michigan State University

ALFRED G. DIETZE

A. INTRODUCTION

This paper compares typal analysis with factor analysis on previously published data by Ekman (4) on the dimensions of emotion. The factor analysis is Ekman's; the typal analysis the writer's.

The comparison has a two-fold purpose. It calls attention to several useful, though little known, analytic techniques developed by McQuitty (cf. references) using data of different nature than employed by the latter to illustrate his methods. It seems desirable to test these methods on a variety of problems and to scrutinize them in regard to their applications and limitations. The data of Ekman's study are ideal for this purpose since the elements of classification are verbal, and one's sense for language helps to appraise the argument. In addition the writer was interested in discovering what light, paralleling or supplementing factor analysis, typal analysis can shed on the relations within sets of data such as herein illustrated.

There is need for objective methods for discovering meaningful categories into which items classify on the basis of measures of association between them. Cattell (1) reviewed the then available methods of cluster analysis and pointed out the desirability of such methods as well as their limitations, e.g., laborious search procedures, lack of exhaustiveness, subjective determination of the limits within which interitem association will be accepted, etc. In more recent years, however, McQuitty (5, 6, 7, 8, 9, 10, 11) came up with a number of ingenious procedures which accomplish this purpose simply, effectively, and economically. We shall refer to these methods collectively as methods of typal analysis since they are all based on a theory of types and have the purpose of revealing the typal structure of a set of data.¹

* Received in the Editorial Office on April 9, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

† This study was supported in part by the All-University Research Fund of Michigan State University.

¹ This usage lumps together various procedures based on typal theory under

The favored method for reducing a matrix of correlation coefficients into its fundamental components has been factor analysis. Although factor analysis has been used on matrices containing other measures of covariation than correlation coefficients, e.g., agreement scores, similarity judgments, etc., the rationale for this practice has not been clearly spelled out. One may question, for instance, whether domains whose classification depends upon the presence or absence of specific characteristics in varying combinations appropriately map into the factor model. A case that comes to mind is the Linnaean system of simple classification which has served biological science well for two and a half centuries. Another is the Kraepelinean classification of mental disorders where symptoms have different values in different syndromes. In the case of similarity measures, too, quantitatively similar judgments of similarity may have different determinants for the perception of similarity. If so, typal classification may be more appropriate to them than classification based on the linear coordinates of factor analysis.

An attractive feature of many typal procedures is economy of time and labor. Desk calculation of a factor analysis is laborious, and with large arrays it becomes prohibitive. Typal analysis accomplishes the reduction of a matrix into types with much less effort and can, therefore, open areas for investigation to many who lack access to electronic computers. McQuitty's methods have, however, been programmed for computer use. Here, being more chary of memory space and permitting easy partialling of the analysis, they place no limits on the size of arrays that can be handled. These considerations should appeal to researchers having access to electronic computation but wishing to deal with more massive amounts of data than can readily be processed by factor analysis. Above economy, of course, theoretical considerations regarding the nature of the data and appropriateness of the model must determine the choice of analytic method.

To place before the reader as much as may be needed to follow the present argument without reference to sources I shall describe the source of the data on which the comparison is based, the results of Ekman's factor analysis, the nature of the typal analytic methods herein employed, and the results derived from typal analysis. A final section summarizes certain conclusions regarding the usefulness of typal analytic methods for the problem illustrated.

the general category of "typal methods." McQuitty, unfortunately, has used factor analytic terminology in naming some of his procedures as, e.g., *elementary factor analysis*. It seems desirable, however, to maintain a distinction in the nomenclature relating to the two methods.

B. SOURCE OF THE DATA

The data of Ekman's study begin with similarity judgments. In a series of papers (2, 3, 4) Ekman described an interesting method of similarity analysis for the investigation of qualitatively different experiences that can be judged comparatively. The experimenter presents all possible pairs of a set of stimuli to subjects pairwise with instructions to judge the subjective similarity of each two stimuli on a five-point scale running from not-at-all-similar to identical. With n stimuli this procedure generates a matrix of similarity scores for each subject. The piled results of all subjects are collapsed by summing over S s and dividing by a factor to express the average relationships on a scale running from 0 to 100. The resulting matrix, consisting of group averages, is finally factored and rotated to simple structure. The extracted factors are interpreted as dimensions of the domain under consideration.

The method is general and can be applied to any set of experience variables for which comparative judgments of similarity can be obtained. Ekman (2, 3) applied it to colors and (4) to words denoting emotions. We have selected for the present comparison his 1955 study of the dimensions of emotion, a study which has served as a model for other work on the expression of emotion, e.g., by Nummenmaa and Kauranne (13), and by Nummenmaa (12).

In this study 23 Swedish words denoting a wide variety of emotional states served as stimuli. S s rated the similarity of the emotional states represented by these words according to the procedure already described. Table 1 reproduces Ekman's matrix in full except that I have rearranged the items in such a way as to group together those falling into the types yielded by typal analysis. Ekman's number designations for the items have been retained.

C. THE FACTOR ANALYSIS: DIMENSIONS OF EMOTION

The purpose of factor analysis is to project the indices of association in a matrix upon a best fitting set of independent, and preferably orthogonal, coordinates. Factor loadings express the degree of projection of the respective variables upon these coordinates. It is based upon a model of k -dimensional space where k is less than n and represents the number of independent factorial dimensions which parsimoniously accounts for the variance in the table. Factors, therefore, represent ideal dimensions of which the variables are imperfect estimates. A strong point made in favor of factor analysis is the accomplishment of classification entirely from the mathematical relation-

TABLE I
MATRIX OF SIMILARITY MEASURES*
(From Ekman, 1955)

| Concept | 6 | 13 | 11 | 10 | 9 | 23 | 21 | 15 | 22 | 17 | 16 | 8 | 4 | 19 | 5 | 14 | 3 | 20 | 12 | 18 | 2 | 1 | 7 |
|----------------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|-----------------|----|----|
| 6-Glad | | 84 | 70 | 43 | 34 | 21 | 33 | 03 | 01 | 05 | 05 | 01 | 00 | 00 | 00 | 00 | 00 | 00 | 11 | 00 | 10 | 00 | 01 |
| 13-Gay | 84 | | 45 | 57 | 19 | 09 | 26 | 03 | 01 | 06 | 03 | 01 | 00 | 00 | 01 | 01 | 00 | 01 | 04 | 00 | 05 | 00 | 00 |
| 11-Happy | 70 | 45 | | 21 | 50 | 39 | 31 | 03 | 02 | 03 | 01 | 00 | 01 | 01 | 00 | 00 | 01 | 00 | 16 | 02 | 10 | 00 | 00 |
| 10-Animated | 43 | 57 | 21 | | 10 | 04 | 12 | 17 | 05 | 29 | 19 | 08 | 04 | 01 | 00 | 03 | 06 | 08 | 05 | 01 | 10 | 02 | 05 |
| 9-Affectionate | 34 | 19 | 50 | 10 | | 80 | 46 | 09 | 10 | 04 | 06 | 01 | 01 | 07 | 03 | 03 | 01 | 01 | 26 | 12 | 24 | 00 | 01 |
| 23-Tenderness | 21 | 09 | 39 | 04 | 80 | | 33 | 11 | 13 | 02 | 02 | 00 | 01 | 13 | 06 | 06 | 01 | 01 | 29 | 18 | 22 | 01 | 00 |
| 21-Benevolent | 33 | 26 | 31 | 12 | 46 | 33 | | 04 | 07 | 01 | 01 | 00 | 01 | 04 | 02 | 01 | 01 | 00 | 07 | 03 | 04 | 00 | 00 |
| 15-Agitated | 03 | 03 | 03 | 17 | 09 | 11 | 04 | | 79 | 65 | 44 | 40 | 33 | 21 | 36 | 37 | 14 | 11 | 19 | 25 | 12 | 06 | 08 |
| 22-Anxious | 01 | 01 | 02 | 05 | 10 | 13 | 07 | 79 | | 44 | 31 | 26 | 46 | 30 | 33 | 41 | 10 | 09 | 17 | 28 | 05 | 06 | 05 |
| 17-Restless | 05 | 06 | 03 | 29 | 04 | 02 | 01 | 65 | 44 | | 64 | 44 | 13 | 15 | 24 | 25 | 15 | 13 | 29 | 20 | 22 | 05 | 09 |
| 16-Impatient | 05 | 03 | 03 | 19 | 06 | 02 | 01 | 44 | 31 | 64 | | 60 | 07 | 09 | 17 | 15 | 33 | 29 | 25 | 17 | 22 | 10 | 11 |
| 8-Irritated | 01 | 01 | 01 | 08 | 01 | 00 | 00 | 40 | 26 | 44 | 60 | | 16 | 09 | 25 | 23 | 55 | 41 | 07 | 07 | 10 | 24 | 27 |
| 4-Frightened | 00 | 00 | 00 | 04 | 01 | 01 | 01 | 33 | 46 | 13 | 07 | 16 | | 09 | 14 | 31 | 15 | 12 | 02 | 06 | 02 | 16 | 09 |
| 19-Sad | 00 | 00 | 01 | 01 | 07 | 13 | 04 | 21 | 30 | 15 | 09 | 09 | 09 | | 75 | 54 | 07 | 12 | 23 | 47 | 05 | 05 | 00 |
| 14-Depressed | 00 | 01 | 01 | 00 | 03 | 06 | 02 | 36 | 33 | 24 | 17 | 25 | 14 | 75 | | 49 | 20 | 15 | 35 | 19 | 06 | 11 | 11 |
| 5-Desperate | 00 | 01 | 00 | 03 | 03 | 06 | 01 | 37 | 41 | 25 | 15 | 23 | 31 | 54 | 49 | | 19 | 19 | 15 | 27 | 10 | 10 | 13 |
| 3-Angry | 00 | 00 | 00 | 06 | 01 | 01 | 01 | 14 | 10 | 15 | 33 | 55 | 15 | 07 | 20 | 19 | | 72 | 03 | 05 | 02 | 26 | 37 |
| 20-Irreful | 00 | 01 | 01 | 08 | 01 | 01 | 00 | 11 | 09 | 13 | 29 | 41 | 12 | 12 | 15 | 19 | 72 | | 02 | 06 | 07 | 43 | 48 |
| 12-Longing | 11 | 04 | 16 | 05 | 26 | 29 | 07 | 19 | 17 | 29 | 25 | 07 | 02 | 23 | 19 | 15 | 03 | 02 | | 65 | 55 | 01 | 03 |
| 18-Want | 00 | 00 | 02 | 01 | 12 | 18 | 03 | 25 | 28 | 20 | 17 | 07 | 06 | 47 | 35 | 27 | 05 | 06 | 65 | | 32 ^a | 01 | 03 |
| 2-Desire | 10 | 05 | 10 | 10 | 24 | 22 | 04 | 12 | 05 | 22 | 22 | 10 | 02 | 05 | 06 | 10 | 02 | 07 | 55 | 32 | | 04 | 09 |
| 1-Disgust | 00 | 00 | 00 | 02 | 00 | 01 | 00 | 06 | 06 | 05 | 10 | 24 | 16 | 05 | 11 | 10 | 25 | 43 | 01 | 01 | 04 | | 61 |
| 7-Rancorous | 01 | 00 | 00 | 03 | 01 | 00 | 00 | 08 | 05 | 09 | 11 | 27 | 09 | 00 | 11 | 13 | 37 | 48 | 03 | 03 | 09 | | 61 |

* The original table has been rearranged to group together items which form types in elementary linkage analysis.

ships among the variables. Yet the process is not free from subjective elements: e.g., the decision as to how many factors to extract, or at what point to stop rotation, has to be made.

1. *Eckman's Factors*

Ekman used Thurstone's centroid method of factor analysis, continuing blind rotation until the distribution of residuals was within the distribution of differences between random halves of his subjects. Table 2 reproduces his rotated factors in an arrangement to reveal the factor structure in the table in terms of highest loadings.

Except for the nature of the coefficients analyzed (similarity scores, not correlation coefficients) the procedure follows conventional lines. The analysis yielded nine meaningful and two ambiguous clusters and their interpretation is indicated by the labels assigned to them. Two of the factors, LONGING (D) and DISGUST (H), are quite pure, having only negligible loadings on other factors. The remaining factors contain items with considerable secondary loadings on some other factor.

There are, however, some curious outcomes, of special interest being the relation between factors PLEASURE (A) and ANIMATION (E). Here *glad*, 6, has its highest loading on factor A, but considerable loading also on factor E, while *gay*, 13, has its highest loading on factor E and considerable representation in factor A. The native English speaker would tend to put together *happy*, *glad*, and *gay* plus, perhaps, *animated* if he were asked to group the words of the experimental list.² Since the experiment was done in Swedish, one might suppose that the words have somewhat different shades of meaning in the two languages which accounts for the occurrence of two factors. However, this is contradicted by reference to the primary data of Table 1 where it appears that the highest similarity score for any pair of concepts in the entire table is the index of 84 mediating between *glad*, 6, and *gay*, 13, which factor analysis allocates to separate dimensions. It seems odd that factor analysis should suppress the most obvious similarity in the matrix by assigning the members of the pair to two factors when the relation of one of them to the marker variable is as weak as it is, i.e., when the relation of *animated*, 10, to *gay*, 13, is only 57 as compared to the relation of 84 between *gay* and *glad*.

² It was the fact that we felt that *happy*, *glad*, *gay* and *animated* belonged together, and that *benevolent* belongs with *affectionate* and *tenderness*, which led us to wonder how typal analysis would classify these items from the information in Ekman's matrix.

TABLE 2
 EKMAN'S FACTORS ORDERED ACCORDING TO FACTOR CLUSTERING*

| Concepts | G Affec- tion | A Pleas- ure | E Anima- tion | D Long- ing | B Discom- fort | Factors F | | C Agita- tion | I Anger | H Dis- gust | K Un- named | J Un- named |
|----------------|---------------------|--------------------|---------------------|-------------------|----------------------|--------------|-----------|---------------------|------------|-------------------|-------------------|-------------------|
| | | | | | | Fear | Agitation | | | | | |
| 23—Tenderness | 722 | —020 | 013 | 046 | 122 | 003 | —031 | —016 | 007 | —001 | —019 | —019 |
| 9—Affectionate | 643 | 059 | 020 | 026 | 037 | —022 | —009 | 005 | —003 | 130 | 037 | 037 |
| 11—Happy | 354 | 629 | 013 | —035 | 010 | —003 | —042 | 018 | —007 | —030 | —012 | —012 |
| 6—Glad | 153 | 626 | 370 | —037 | 018 | —002 | —020 | 012 | 000 | 011 | —015 | —015 |
| 13—Gay | —019 | 373 | 626 | —030 | 048 | —001 | 013 | 008 | 003 | 070 | —008 | —008 |
| 10—Animated | 024 | 017 | 678 | 014 | —024 | —003 | 223 | 052 | 010 | —011 | 071 | 071 |
| 12—Longing | 074 | 089 | —033 | 733 | 200 | 021 | 005 | 004 | —028 | 018 | 118 | 118 |
| 2—Desire | 045 | —003 | 025 | 548 | —042 | —032 | 000 | 015 | 036 | 061 | 366 | 366 |
| 18—Want | —031 | —025 | 007 | 630 | 482 | 079 | —003 | —010 | —029 | 037 | 028 | 028 |
| 5—Desperate | —003 | 000 | 034 | —007 | 464 | 271 | 139 | 119 | 007 | 015 | 427 | 427 |
| 14—Depressed | —009 | 036 | 024 | —053 | 758 | 021 | 216 | 003 | 035 | —006 | 292 | 292 |
| 19—Sad | 032 | —023 | 078 | 030 | 859 | —030 | 110 | —005 | 015 | —015 | 251 | 251 |
| 22—Anxious | 056 | 002 | —024 | 046 | 218 | 744 | 417 | 008 | 023 | 079 | 046 | 046 |
| 4—Frightened | —004 | —008 | —008 | —017 | —003 | 594 | 026 | 090 | 010 | 056 | 235 | 235 |
| 15—Agitated | 081 | 000 | 085 | 024 | 181 | 567 | 677 | 018 | 044 | 008 | —013 | —013 |
| 17—Restless | —009 | —028 | 197 | 129 | 040 | 193 | 717 | 109 | 024 | 007 | 131 | 131 |
| 16—Impatient | —030 | 028 | 058 | 153 | —040 | —010 | 681 | 357 | —019 | 033 | 056 | 056 |
| 8—Irritated | —008 | 050 | —020 | —023 | —026 | 062 | 510 | 493 | 047 | 005 | 114 | 114 |
| 3—Angry | 018 | —025 | 021 | —004 | —017 | 005 | 097 | 768 | —001 | —003 | 009 | 009 |
| 20—Ireful | 017 | —030 | 053 | 050 | —010 | 004 | 003 | 594 | 215 | —019 | —015 | —015 |
| 1—Disgust | 000 | 004 | —005 | 010 | —018 | 045 | —021 | —022 | 681 | 003 | 008 | 008 |
| 7—Rancorous | —025 | 005 | —010 | 048 | —009 | —004 | —019 | 046 | 670 | 026 | 006 | 006 |
| 21—Benevolent | 040 | 016 | —006 | —006 | 012 | —013 | 019 | 003 | 000 | 507 | —006 | —006 |

* The column headings are Ekman's factor labels. The capital letters above them designate their order of extraction.

Remarkable also is the isolation of *benevolent*, 21, in factor K with no appreciable loadings on any other factor. *Benevolent* is semantically no more unique than a number of other items in the list. It would seem to be related to *affectionate* and *tenderness*, a humanitarian wanting the welfare of others. Table 1 does, indeed, show *benevolent* to be most similar to *affectionate*, 9, and then to *tenderness*, 23, the indices of similarity being sufficiently substantial to support classification of the three emotions in the same cluster even though the mathematical operations of factor analysis do not do so.

D. THE TYPAL ANALYSIS: TYPES OF EMOTIONS

The concept of type is applied here only to items for which measures of association have been entered in a matrix. Whether types found in such matrices reflect the universe of relations from which they are presumably drawn depends upon sampling considerations as does any statistical inference. Types, as well as dimensions, are products of man's ordering that which in nature is unordered, for, as Lamarck has said, "The classifications are artificial, for nature has created neither classes nor orders nor families nor kinds of permanent species, but only individuals."

The structure of nature is, accordingly, a matter of variety. However, if we consider a collection of individuals and compare them in regard to attributes which they appear to us to share, we end up with groupings that aid comprehension. Taking any set of such items, we can define a type with reference to that set as a subset of such nature that each member of the subset is more like all other members of the subset than it is like any individual outside of that subset. The procedure which McQuitty (10, 11) calls "typal analysis"³ proceeds from such a definition. It yields at least two dyadic types from any matrix which contains types. The elements of such dyads are always reciprocal, each member of the pair being more like the other than it is like any other item in the matrix. Larger types consist of additional members which are more like the dyadic pair and each other than they are like items outside the type. Since types may coalesce at higher levels the method may also be used to reveal the hierarchical typal structure contained in a matrix. The above definition will be referred to as definition 1 in subsequent sections of the paper.

Due to errors of observation (and "errors" of nature) empirical data

³ Not to be confused with the generic meaning of typal analysis as defined for this paper. "Typal analysis" is one of McQuitty's techniques; typal analysis, generically, is any method which classifies the elements of a matrix into types on the basis of measures of association between them, factor analysis excepted.

frequently contain impure types as well as pure types. In that event a less restrictive definition of type may be desirable. This definition, hereinafter referred to as definition 2, states that a type is a subset of a larger set of items whose members are more like *some* other member of the subset than they are like any item outside of the subset. This definition underlies McQuitty's elementary linkage analysis (1957) which is the base from which the present analysis proceeds. Types which satisfy this definition usually contain within them types which satisfy also the first definition, or may be included within more inclusive types which satisfy either definition, or both. The less stringent definition in this paragraph will be referred to as definition 2.

In the context of this paper a typal method of analysis will be considered to be any procedure which reveals types in a set of data indexing relations between all pairs of items in the set under study. Typal analysis may proceed to any level of generality: it may be used to uncover all elementary relationships which satisfy the definition of type selected; or it may be used to reveal the complete hierarchical typal structure implicit in the set of data. Suitable for typal analysis are all matrices tabling acceptable measures of relationship between pairs of items, e.g., correlation coefficients, agreement scores, percentage overlap, similarity measures, D^2 , etc.

1. *Elementary Linkage Analysis Applied to Ekman's Data*

McQuitty's several methods accomplish classification in different ways, for different purposes, and with varying degrees of completeness of the information they yield. The present analysis uses basically elementary linkage analysis (6), but borrows also from other typal analytic techniques described by that author and introduces some innovations of the present writer which seem consistent with typal theory.

a. Method. The procedure of elementary linkage analysis is simple. (a) The coefficients of association between items are entered into a matrix as is done in factor analysis. (b) The highest entry in every column is marked, thus identifying the item in the rows which is most like the item heading the column. (c) The highest entry in the entire table is now identified from among the marked items. Since this involves a reciprocal pair of items, this value will occur in two columns at the intersects marked by the reciprocal items. Such reciprocal pairs are pure types in the sense of definition 1. The first such pair isolated in the matrix, i.e., the pair identified by the highest coefficient in the table, constitutes the axis of the first type and the item symbols are written in a diagram representing the type with two arrows

pointing in opposite directions between them to show that they are reciprocal (see below). (*d*) The two rows which pass through the intersect with these columns are now searched for marked entries which identify items that have their highest relations to each member of the axial pair. Items so identified are added to the type by means of arrows showing the direction of highest relation. Rows intersecting the columns so found are similarly searched until no further items joining the type are found. (*e*) The highest remaining entry in the table is now identified, constituting the axis of a second type, and steps *d* and *e* are repeated. This is continued until all items in the matrix have been assigned to some type. Columns and rows involving items joining types are, of course, eliminated from further consideration as soon as they have been used in accordance with the above rules.

This procedure will yield at least one type for every matrix, although an outcome of only one type would lead to the rejection of a typal hypothesis for the data concerned. The diagrams of typal linkages serve as excellent maps of typal structures in accordance with definition 2, showing for each item the other item with which it has its highest relation.⁴

CHART 1
TYPAL LINKAGES OF EKMAN'S EMOTIONS

| Label | Linkage diagram | Key to Items |
|----------------|---|---|
| Type A. JOY | $(6) \Rightarrow (13)$ $\uparrow \quad \uparrow$ $(11) \quad (10)$ | 6. Glad 13. Gay 11. Happy 10. Animated |
| Type B. LOVE | $(9) \Rightarrow (23)$ \uparrow (21) | 9. Affectionate 23. Tenderness 21. Benevolent |
| Type C. UNREST | $(15) \Rightarrow (22)$ $\uparrow \quad \uparrow$ $(17) \quad (4)$ \uparrow (16) \uparrow (8) | 15. Agitated 22. Anxious 17. Restless 4. Frightened 16. Impatient 8. Irritated |

⁴ Any other relationship satisfying definition 1, e.g., second-most-like, third-

CHART 1 (continued)

| | | |
|--------------------|----------------------------------|--|
| Type D. DISCOMFORT | (14) \rightleftharpoons (19) | 14. Depressed 19. Sad 5. Desperate |
| | ↑ (5) | |
| Type E. ANGER | (3) \rightleftharpoons (20) | 3. Angry 20. Ireful |
| Type F. LONGING | (12) \rightleftharpoons (18) | 12. Longing 18. Want 2. Desire |
| | ↑ (2) | |
| Type G. DISGUST | (1) \rightleftharpoons (7) | 1. Disgust 7. Rancorous |

b. Linkage types from Ekman's data. The linkage types yielded by Ekman's data are as follows (Chart 1). Where types are identical with factors, Ekman's factor designations have been used to label them; where they differ from factors, suitable new labels are provided to distinguish them.

Of seven types realized, four correspond to factors: type D = factor B, DISCOMFORT; type E = factor I, ANGER; type F = factor D, LONGING; and type G = factor H, DISGUST. DISCOMFORT, ANGER, and DISGUST constitute pure types, each member of the type being more like all $n - 1$ other members of the type than it is like any element outside the type (type definition 1). Since this is always true of reciprocal pairs which form the axes of types, such reciprocal pairs may be considered, if taken by themselves, as pure types, thus providing an objective basis for interpreting and naming the type, for there are always as many reciprocal pairs in a matrix as types which may be extracted from it. DISCOMFORT has three elements conforming to the definition of a pure type; but LONGING falls short of this definition since the element, *want*, has its second highest similarity with the element *sad* in type DISCOMFORT. This finding is, however, instructive. It points to the fact that verbal stimuli do not have invariant meanings: different contexts give them different values in much the same way as the judgment of "heavy" may have different values when it occurs in a light series than when it occurs in a heavy series. Wanting in juxtaposition with longing or desire is seen as

most-like, etc., can be shown by means of suitably identified arrows drawn between items. For more than four elements it would be necessary to arrange the elements around the circumference of a circle. A type consisting of n items will have a maximum of $n(n-1)$ such relationships.

similar to the latter; in juxtaposition with sad the relationship between want and sadness is elicited: to be in want is to be sad, indeed. Typal analysis allows for an indicant to have different values in different situations, which factor analysis does not. Thus it may direct attention to the source of ambiguities and to the nature of intertype relationships.

The three remaining types contain larger numbers of elements than the factors. Type JOY combines factors AFFECTION and ANIMATION. This makes good sense since it preserves the strongest reciprocal relationship between any two elements in the entire matrix, between *glad* and *gay* which factor analysis relegated to two separate factors. Actually JOY is a fairly tightly structured category, as will be seen from the following considerations. For any element in a type of n elements there are $n - 1$ relationships testing type definition 1: most like, second most like, etc. For the entire subset there are $n(n - 1)$ such relations. Thus, for a type containing four elements one may make $4(4 - 1) = 12$ such tests. Reference to the columns of Table 1 in which the elements of type JOY appear will reveal that nine such tests are satisfied: four relations of most-like; three of second-most-like; and two of third-most-like. *Happy* has its second-most-like relation to *affectionate* and its third-most-like relation to *benevolent*, both in the LOVE cluster; and *animated* has its third-most-like relation to *restless* in the UNREST cluster. And is it not the case that when one feels affectionate or benevolent one is also apt to be happy? Thus the juxtaposition of affection and happiness in experience may cause them to be confused, or seen as similar, in perception. The relation of animation to restlessness as well as to joy may be due to similar considerations, but I am more inclined to think that it may be an attribute of various emotional states in the sense of level of energy release. This seems to be the case with happy, glad, and gay involving increasing degrees of animation on a continuum from repose to restlessness.

In type LOVE *benevolent* is tied in with the elements of factor AFFECTION. Benevolence, indeed, involves a tender feeling for others less fortunate than the person experiencing benevolence. Thus an element which factor analysis left isolated is brought into association with a type.

Type UNREST incorporates factors AGITATION and FEAR. It is a more loosely organized cluster than the others, although of the $n(n - 1) = 30$ relationships testing definition 1, 21 are found in the cluster. The unifying concept in the cluster seems to be *agitated*, which suggests the label UNREST. If one considers *agitation* as a quality of emotions, one may form a submatrix of the remaining elements of this cluster and subject it to elementary linkage analysis. This yields Ekman's factors ANGER and FEAR under agitation as a

superordinate. In our chart of the hierarchical typal structure of Ekman's matrix (Chart 2), we have shown it thus broken down. Both anger and fear are highly agitated states.

c. *Summary.* In the light of the above discussion the following summary comments may be made:

(1). Elementary linkage analysis reduces the original 23 elements of Ekman's matrix to seven types, while factor analysis extracts ten plus factors. Elementary linkage analysis seems, therefore, to accomplish a more thorough reduction of the data into classes whose elements are meaningfully related.

(2). Elementary linkage analysis incorporates all elements of a matrix into some type in which it best fits. Factor analysis carried out to some specified number of rotations may leave elements unassociated. In the present case *benevolent*, left isolated by factor analysis, is seen as meaningfully related to *affection* and *tenderness*. However, elementary linkage analysis cannot isolate an element so that it stands by itself. Thus it may not be able to reveal elements that are really foreign to the rest of the set, unless its relations to all other elements of the set is zero.

(3). Elementary linkage analysis permits direct objective observation of the actuality of the types found by comparison with the order of relationships indexed in the matrix. In fact, the method unfailingly yields all most-like relationships, leaving for further inspection, if desired, only the next-most relationships that need to be considered, to test definition 1.

(4). Elementary linkage analysis allows pinpointing other structural relationships within types and between types. One may, in a sense, observe how a type shades off into not-type, or into related type. Also, since elements join types in descending order of the magnitudes of the measures of association between them, the order of concatenation of the elements may be easily observed in the type diagrams.

2. *Typal Loadings*

Elementary linkage analysis as so far described reveals types and their basic structure. It does not, however, quantify the degree of relationship which elements bear to types in the manner of factor loadings indexing the projection of variables on the coordinates of the factorial space. McQuitty (6), however, showed that typal relevancies may be computed from submatrices of the elements of types taken separately. First factor loadings for each such submatrix are calculated in the usual manner, and these approximate rotated factor loadings when types correspond to factors (6). McQuitty (10, 11) described still other methods for estimating loadings on types.

Unfortunately in naming these procedures he used the label "elementary factor analysis" which fails to distinguish them from factor analysis. However, these methods have simplicity and straightforwardness in their objective reference to the basic data in the matrix to recommend them.

One of these methods selects "prototypes," i.e., the elements which are most characteristic of the several types found. Submatrices containing only the elements of a single type are formed for each type, and columns are summed. The column yielding the highest sum is the prototype of the element composing the type, and the entries in the corresponding columns of the primary matrix are the loadings of the several variables. For the present analysis *glad*, *affectionate*, *agitated*, *sad*, *angry*, *longing*, and *disgust* are the prototypes of types A to G respectively and the entries under them in Table 1 are the loadings.

The writer feels that more representative loadings may be calculated in terms of the mean index of association of each item with the elements composing the respective types. Such an index should be less subject to chance fluctuation. In the present connection where similarity measures are used such an index would have the meaning of mean similarity with the type. If the matrix is arranged so that types are segregated into blocks, as in Table 1, it is a simple matter to sum across rows in each block and divide by the number of columns composing the respective types. Diagonals may be filled with 1.00s, or with the highest entry in the row. We have followed the second procedure to obtain the values for Table 3, assuming the similarity index of an item with a member of a type to be an estimate of its similarity with the composite type, and the reliability of an element to be the highest similarity it has with any other member of the type.⁵

3. Hierarchical Typal Structure

McQuitty (7, 8, 9) described how one may take the intersects of columns and rows containing prototypes as measures of relation between types, form them into a matrix tabling the relations between prototypes, and subject the submatrix to elementary linkage analysis to discover higher order types, a process which may be repeated for higher order types until no further linking

⁵ Thus the typal loadings of the first three items in Type 1 were calculated as follows:

| Element | 6 | 13 | 11 | 10 | Mean of Row |
|---------|----|----|----|----|----------------|
| 6 | 84 | 84 | 70 | 43 | 70 |
| 13 | 84 | 84 | 45 | 57 | 68 |
| 11 | 77 | 45 | 70 | 21 | 52 |

TABLE 3
TYPAL LOADINGS*

| Concept | Type A JOY | Type B LOVE | Type C UNREST | Type D DISCOMFORT | Type E ANGER | Type F LONGING | Type G DISGUST |
|----------------|---------------|----------------|------------------|----------------------|-----------------|-------------------|-------------------|
| 6-Glad | 70 | 29 | 02 | 00 | 00 | 07 | 00 |
| 13-Gay | 68 | 18 | 02 | 01 | 00 | 03 | 00 |
| 11-Happy | 52 | 40 | 02 | 01 | 00 | 09 | 00 |
| 10-Animated | 44 | 09 | 14 | 01 | 07 | 05 | 04 |
| 9-Affectionate | 28 | 69 | 05 | 04 | 01 | 21 | 00 |
| 23-Tenderness | 18 | 63 | 05 | 08 | 01 | 23 | 00 |
| 21-Benevolent | 26 | 42 | 02 | 02 | 00 | 05 | 00 |
| 15-Agitated | 06 | 08 | 57 | 31 | 12 | 19 | 07 |
| 22-Anxious | 02 | 10 | 51 | 35 | 10 | 17 | 06 |
| 17-Restless | 11 | 02 | 49 | 21 | 14 | 24 | 07 |
| 16-Impatient | 08 | 03 | 45 | 14 | 31 | 21 | 10 |
| 8-Irritated | 03 | 00 | 41 | 19 | 48 | 08 | 26 |
| 4-Frightened | 01 | 01 | 27 | 18 | 14 | 03 | 12 |
| 19-Sad | 00 | 08 | 16 | 68 | 10 | 25 | 02 |
| 14-Depressed | 00 | 04 | 25 | 66 | 18 | 20 | 11 |
| 5-Desperate | 01 | 03 | 29 | 52 | 19 | 17 | 12 |
| 3-Angry | 02 | 01 | 24 | 15 | 72 | 05 | 32 |
| 20-Ireful | 02 | 01 | 19 | 15 | 72 | 08 | 46 |
| 12-Longing | 09 | 21 | 16 | 19 | 02 | 62 | 02 |
| 18-Want | 01 | 11 | 17 | 36 | 06 | 54 | 02 |
| 2-Desire | 09 | 17 | 12 | 07 | 05 | 47 | 06 |
| 1-Disgust | 00 | 00 | 11 | 09 | 34 | 02 | 61 |
| 7-Rancorous | 01 | 00 | 12 | 08 | 42 | 05 | 61 |

* The order of the elements corresponds to that of Table 1. It places together the elements belonging to each type.

occurs. Instead of following this procedure, we have calculated intertypal similarity indices by collapsing the columns associated with types in Table 3, yielding the mean intertype similarities of Table 4.

TABLE 4
MEAN INTERTYPE SIMILARITIES

| Type | A | B | C | D | E | F | G |
|------|----|----|----|----|----|----|----|
| A | | 24 | 05 | 01 | 02 | 06 | 01 |
| B | 24 | | 04 | 05 | 01 | 16 | 00 |
| C | 05 | 04 | | 23 | 21 | 15 | 11 |
| D | 01 | 05 | 23 | | 15 | 21 | 08 |
| E | 02 | 01 | 21 | 15 | | 04 | 38 |
| F | 06 | 16 | 15 | 21 | 04 | | 04 |
| G | 00 | 00 | 11 | 08 | 38 | 04 | |

Elementary linkage analysis applied to this table yields three higher order types as follows:

Higher order type

Label

Elements

I

REJECTION

Types E and G

II

ACCEPTANCE

Types A and B

III

DISTURBANCE

Types C, D, and F

Assigning superordinate labels to these we may form a chart of the hierarchical typal structure as in Chart 2.

CHART 2
HIERARCHICAL TYPAL STRUCTURE OF EKMAN'S SIMILARITY MATRIX

| EMOTIONS | | | | | | |
|------------|--------------|-----------|-----------|-------------|------------|---------|
| ACCEPTANCE | | REJECTION | | DISTURBANCE | | |
| JOY | LOVE | ANGER | DISGUST | UNREST | DISCOMFORT | LONGING |
| happy | affectionate | angry | disgust | agitated | sad | longing |
| glad | tenderness | ireful | rancorous | restless | depressed | want |
| gay | benevolent | | | impatient | desperate | desire |
| animated | | | | irritated | | |

As may now be seen, the 23 elements of Ekman's matrix fall into seven clusters at the typal level which we have labelled JOY, LOVE, ANGER, DISGUST, UNREST, DISCOMFORT, and LONGING. These, in turn, combine to form the higher classes, ACCEPTANCE, REJECTION, and DISTURBANCE. Joy and love involve an accepting attitude toward their object, hence ACCEPTANCE; anger and disgust are rejecting in nature; and unrest, discomfort, and longing may be considered disturbed states. The items under UNREST are arranged according to the breakdown suggested on page 154, giving an *impatience* subcluster and a *fear* subcluster under *agitated*. Thus the operations of elementary linkage analysis and accessory procedures provide a meaningful diagram of the relations between the emotions judged in Ekman's experiment.

E. SUMMARY AND CONCLUSIONS⁶

We have analyzed data by Ekman on the dimensions of emotion by a typal analytic procedure developed by McQuitty in order to illustrate the latter method on different data than employed by McQuitty, and in order to compare the results of such an analysis with those of factor analysis. We find the following:

1. Elementary linkage analysis yields types comparable with, if not superior to, factor analysis in regard to the psychological sense they seem to make.
2. Elementary linkage analysis accomplishes its objective more economically than factor analysis. The present matrix containing 23 rows and columns

⁶ Wrigley (1957) has examined McQuitty's (1955) method of pattern analysis, also comparing it with factor analysis.

can be analyzed in about ten minutes as compared with hours of labor which hand computation of factors would require.

3. Linkage diagrams provide readily understandable pictures of the main relations between elements of types. Reference to the primary matrix makes it possible to supplement these diagrams to show all relations which conform to a definition of a perfect type.

4. Elementary linkage analysis accomplishes a more thorough reduction of the matrix than factor analysis. Whether this is always the case requires further experience with the method.

5. Several methods for quantifying the relationship of elements to types are available: typal relevancies; loadings on prototypes, mean typal similarities. Typal plots may be drawn utilizing these data in the same way that factor plots are constructed.⁷

6. Typal loadings may be further collapsed to yield measures of intertype relationships. These may be entered into matrices and subjected to elementary linkage analysis to discover higher order classes, a process which may be continued until the data no longer yield superordinate classes. This yields a picture of the hierarchical structure of the data being considered.

7. At all points of typal analysis the results can be checked against the primary data for conformity to the definition of type.

8. Typal analysis is, of course, appropriate where typal theory holds, or may be assumed to hold. It may be that some natural happenings are best understood within a typal framework, and that others conform to a coordinate factor model. It may also be that the two models supplement each other. The author feels that the typal model better fits Ekman's elements than the factor model.

REFERENCES

1. CATTELL, R. B. A note on correlation clusters and cluster search methods. *Psychometrika*, 1944, **9**, 169-184.
2. EKMAN, G. Dimensions of color vision. *J. of Psychol.*, 1954, **38**, 467-474.
3. ———. Eine neue Methode zur Erlebnissanalyse. *Z. exper. u. angew. Psychol.*, 1954, **2**, 167-174.
4. ———. Dimensions of emotion. *Acta Psychol.*, 1955, **11**, 279-288.
5. MCQUITTY, L. L. Agreement analysis: Classifying persons by predominant patterns of response. *Brit. J. Stat. Psychol.*, 1956, **9**, 5-16.
6. ———. Elementary linkage analysis for isolating orthogonal and oblique types and typal relevancies. *Educ. & Psychol. Meas.*, 1957, **17**, 207-229.

⁷ Typal plots are not included in this article on account of space considerations. However, the effort to draw them will repay the reader in bringing out the fact that our linkage types are orthogonally related.

7. ———. Hierarchical linkage analysis for the isolation of types. *Educ. & Psychol. Meas.*, 1960, **20**, 55-67.
8. ———. Hierarchical syndrome analysis. *Educ. & Psychol. Meas.*, 1960, **20**, 293-294.
9. ———. Comprehensive hierarchical analysis. *Educ. & Psychol. Meas.*, 1960, **20**, 55-67.
10. ———. Elementary factor analysis. *Psychol. Rep.*, 1961, **9**, 71-78.
11. ———. Typal analysis. *Psychol. & Educ. Meas.*, 1961, **21**, 677-696.
12. NUMMENMAA, T. Dimensions of verbal and facial expressions of emotion as derived from matched stimuli. *Acta Acad. Paedagog. Yväs kylä.*, 1960, **20**, 44-42.
13. NUMMENMAA, T., & KAURANNE, U. Dimensions of facial expression. *Rep. Dept. Psychol., Inst. Pedagog., Yväs kylä*, 1958, No. 20.
14. WRIGLEY, C. F. Cluster analysis or factor analysis? The divisional structure of the American Psychological Association. *Psychol. Rep.*, 1957, **3**, 497-506.

Department of Psychology
Michigan State University
East Lansing, Michigan



INFORMATION "CHUNKING" AND SHORT-TERM RETENTION*¹

University of Southern California, Los Angeles

M. S. MAYZNER² AND R. F. GABRIEL

A. INTRODUCTION

Over the past few years many large-scale computerized man-machine systems have been developed, and with this development a variety of new problems have arisen concerning man's capacity to process quickly and accurately large quantities of information. One critical component of such information processing capacity is that of information storage and retrieval or more specifically, man's capacity for retaining and producing at some later time particular items of information.

The present study is designed to examine the effects of one parameter believed to be important to such information storage and retrieval processes in man, namely the effects of information organization or "chunking." Specifically Ss are presented with a display of information items (e.g., digits) for a given period of time and then asked for their immediate recall. Different groups of Ss receive the same items, but organized or "chunked" in different ways and the question asked is: "Do differences in organization significantly affect recall?"

B. METHOD

1. Subjects

The Ss were 50 students from education and psychology classes at the University of Southern California.

2. Stimulus Material and Procedure

Five 12 digit sequences were selected from a table of random numbers under the following constraints. All digits from one through nine were included at least once in each sequence, zeros were excluded, no digit occurred in any sequence more than twice, and any sequence in which a digit was followed by itself was excluded. Each sequence was placed, using 1/4 inch

* Received in the Editorial Office on April 10, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ This research was supported by Contract Nonr 228(24) between the Office of Naval Research and the University of Southern California.

² Now on the faculty of New York University.

high "Letraset" sheet No. 129 numbers, on plain white 5×8 index cards, one sequence per card.

Five different organizations of the 12 digit sequences were prepared. Group 1 received an organization in which each of the 12 digits was placed on a separate line making a vertical column of 12 digits (i.e., 12 lines). Group 2 received an organization in which two digits appeared on each line (i.e., six lines). Group 3 received an organization in which three digits appeared on each line (i.e., four lines). Group 4 received an organization in which four digits appeared on each line (i.e., three lines). Group 5 received an organization in which six digits appeared on each line (i.e., two lines).

There were 10 *Ss* per group and *Ss* were assigned randomly to the several groups. Within each group *S* received five trials on one of the five organizations. On each trial, *S* was presented with the display card for six seconds and then asked for his immediate free recall. The *S* was allowed as much time as he required for his recall before the next trial, and all testing was performed individually.

C. RESULTS AND DISCUSSION

The number of digits recalled correctly in their proper position in a sequence was selected as the dependent response measure. Since each *S* received five trials, the mean of these five trial measures was selected as the criterion measure for each *S*, and Table 1 presents the mean number of digits correctly recalled for the five different organizations, summed across trials.

TABLE 1
MEAN NUMBER OF DIGITS RECALLED FOR THE DIFFERENT ORGANIZATIONAL SEQUENCES

| Organizational sequence | Mean | <i>SD</i> |
|----------------------------|------|-----------|
| Group 1 (12 lines) | 5.34 | 1.95 |
| Group 2 (6 lines) | 6.56 | 2.02 |
| Group 3 (4 lines) | 6.16 | 2.69 |
| Group 4 (3 lines) | 6.80 | 2.38 |
| Group 5 (2 lines) | 7.10 | 2.27 |
| Group 6 (1 line) | 6.28 | 2.06 |

Note: Group 6 tested at a later date than Groups 1-5.

Since Bartlett's test for homogeneity of variance indicated the variances between conditions were not significantly different ($P > .10$), the data were analyzed utilizing Edward's "Repeated Measurements of the Same Subjects" design (1). A significant main effect ($F = 3.14$, $df = 4, 45$; $p < .05$) was obtained for the organizational variable, but no significant

differences were obtained for the trials variable or the organizational by trials interaction term.

The organizational variable appears then to produce a systematic effect on retention, in that over the range examined from 12 lines to two lines, there is an increase in the number of digits retained, with one minor reversal occurring in Group 3. Since a systematic increase was found from 12 lines to 2 lines, it was decided to run an additional group of 10 Ss, who received the same treatment as the previous Ss, except all 12 digits appeared on one line. The result for this group is also included in Table 1 and yielded a mean of 6.28 digits correctly recalled. Unfortunately, as the group was run at a somewhat later time, a different subject population may be the factor accounting for the sudden drop in recall. However, these same 10 Ss participated in another immediate memory study and their performance here replicated results found with the 50 Ss tested earlier, so there is moderately good reason to suppose that their results reflect the effect of the organizational variable and not an S population variable. If this assumption is warranted, then it would appear that somewhere between six and 12 digits per line, there is a decrement in recall performance.

The results of this experiment are supported by a study by Mackworth (2). Specifically, Mackworth varied the number of digits shown simultaneously as well as presentation time and found that retention improved as the number of digits displayed simultaneously increased, particularly at the faster presentation rates. Mackworth cited visual factors as the most likely reason for her results. However, her procedure was somewhat different from the one used in this study, in that the stimuli were presented moving across a display surface, in differing amounts and at different presentation rates.

It is suggested that in the present study the effect of organization may best be handled by the concept of "chunking" developed by George Miller (3). Miller notes that there are two ways in which man can overcome the limitations of immediate memory. The first is that of recoding. The second is chunking. Recoding involves a transformation of information, such as going from the binary to the octal system of numbers, whereas, according to Miller, chunking seems to involve organizing or grouping items and then storing the group. Very little specific information is available concerning the variables involved in chunking. In referring to an experiment in which a memory span of five words was obtained, Miller states that it could just as easily be called a memory span of 15 phonemes, since there were about 3 phonemes per word. Thus, Miller suggests it is the number of chunks that is determined by the memory span and not the amount of information

per chunk. That is, with 2, 3, 4, or 6 numbers per line, there is the possibility of treating the whole line as one or two "chunks" and in this sense the information is first recoded and then retained as a "chunk." When, however, a single line of 12 digits is presented, this recoding and chunking mechanism does not function as well as for more lines with fewer stimuli per line.

In general, the results of the present study certainly suggest that short-term retention is a function in part, at least, of the organization of the material, when all other factors are held constant. Any display design then, in which information storage and retrieval may be a factor, should take such organizational variables into account. Further, it is suggested that a convenient way of conceptualizing the effects of such organizational variables is Miller's "chunking" model of information processing.

D. SUMMARY

The present study was designed to examine the effects of information organization or "chunking" on short-term retention. Five different groups of 10 Ss each received the same set of 12 digits for a period of six seconds, but each group received the 12 digits in a different organizational array, i.e., 12 lines (1 digit per line), six lines (two digits per line), etc. The results showed a systematic increase in retention from 12 lines to two lines and a "chunking" model was employed to account for these effects of stimulus organization on retention.

REFERENCES

1. EDWARDS, A. L. *Experimental Design in Psychological Research*. New York: Rinehart, 1950.
2. MACKWORTH, J. F. The effect of display time upon the recall of digits. *Can. J. Psychol.*, 1962, **16**, 48-55.
3. MILLER, G. A. The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.*, 1956, **63**, 81-97.

15 Sydney Avenue
Deal, New Jersey

THE EFFECT OF FREQUENCY OF PRESENTATION ON THE CHOICE OF NONSENSE SYLLABLES*

E. I. du Pont de Nemours & Company, University of Hawaii, and Princeton University

JAMES C. BECKNELL, JR., WARNER R. WILSON, AND J. C. BAIRD

A. INTRODUCTION

The experiment reported in this paper was designed to investigate a determinant of choice which has received relatively little attention. The basic experimental hypothesis is that choice is affected by frequency of presentation, that is, Ss prefer stimuli they have seen frequently.

It seems well established experimentally that stimulus familiarity is an important determiner of recall (1, 6, 14) of stimulus recognition and response bias [reviewed by Goldiamond (4)]. Although these studies are indirectly relevant to studies of affectivity, the dependent variables used have not included Ss' expressed *preference* for a particular stimulus. Very few investigators have tried to determine the effect of familiarity on preferences. Studies which have been conducted indicate that a person's tastes depend to a certain extent upon familiarity but that other factors may be equally important.

Krugman (9) studied the changes of student preferences for two categories of music, classical and swing, and found that increasing Ss' exposure to music originally disliked increased their preference for that music. Maslow (10), however, has shown that increased familiarity with objects does not necessarily result in increased liking but rather that preferences depend upon several factors, one of which is the type of stimuli used. Hartley (5) also claims that preference shifts are partly a function of the stimulus category. Using different types of paintings Hartley isolated a number of variables which affected an individual's tastes but the magnitude of these changes depended upon the type of stimulus used.

Although the Maslow, Krugman, and Hartley investigations obtained measures of preference and not merely of recall or response bias, their results cannot be clearly attributed to presentation frequency since stimulus meaningfulness and pre-experiment stimulus familiarity were not controlled. It has been previously pointed out by Johnson, Thomson and Frincke (7)

* Received in the Editorial Office on April 17, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

and by Wilson and Becknell (15) that association value of nonsense syllables may influence Ss' preference for such syllables. It would seem that the same effect would be obtained with more meaningful material, as was employed in the earlier studies of preference behavior. One method, whereby the meaningfulness of test stimuli and Ss' prior experience with these stimuli can be controlled, is to use nonsense syllables with known association value (e.g., 3). Johnson *et al.* (7) have employed such a technique and report that increasing the exposure frequency of nonsense words raises their "goodness" as evidenced by semantic differential scores. The generality of these results has been questioned by Stricker (13) who claims that word "goodness" is not equivalent to word "value" which involves some motivational predisposition apart from sheer pleasantness. The distinction between *value* and *goodness* may have important implications for studies of perceptual defense and related concepts [though see rejoinder (8)]. The present experiment took an approach similar to that of Johnson *et al.* (7) to test the hypothesis that Ss prefer nonsense syllables they have seen frequently over syllables they have seen less frequently. However, instead of the semantic differential, the present study used both paired comparison and ranking procedures to obtain indicants of Ss' stimulus preferences.

B. METHOD

1. Subjects

Four different groups served as Ss. Groups 1 and 2, containing 32 and 25 Ss respectively, were members of a women's church group. Groups 3 and 4, containing 34 and 36 Ss respectively, were members of a women's club. The organization participating received one dollar for each of its members' services.

2. Stimuli and Apparatus

Two sets of test stimuli, four nonsense syllables in a set, were used. The stimuli were mounted on slides and presented with a 500 watt slide projector onto a beaded glass screen. The syllables of each set varied somewhat in association value as determined by Glaze (3).¹ The four syllables of Set 1 were B I K, L U F, and N I M with an association value of 67 per cent, and G A H with an association value of 0 per cent. The four syllables of Set 2 were C I V, F I Y, and R E Q with an

¹ A more recent study by Archer (1960) on the association value of nonsense syllables was unpublished at the time the present study was conducted.

association value of 67 per cent, and T O V with an association value of 0 per cent. In addition to the nonsense syllables, slides of eight dummy advertisements and thirty-two landscape scenes were used. Duplicates were made of the slides in order to facilitate repeated presentation of the same stimuli. The dummy ads were used in relation to another experiment, the results of which are not discussed in this paper. The purpose of the landscape pictures was to make the experiment less boring for the Ss.

The test sessions of Groups 1 and 2 were conducted in a small auditorium. The test sessions of Groups 3 and 4 were conducted in a large room.

3. Procedure

The Ss were systematically exposed to nonsense syllables at frequencies of one, four, seven, or ten times during the experimental session. Each syllable was shown to each subject group at a different frequency and within each group all four frequencies were represented, one syllable from each of the two sets. Table 1 shows the experimental design. Interspersed with the nonsense syllables were the dummy advertisements and landscape scenes. The order of presentation was systematically arranged so that repetition of the same syllable would be spaced as widely apart as possible as determined by the syllable's frequency, the dummy ads and the landscapes. First, twenty-two nonsense syllables (Set 1), twenty-two dummy ads and sixteen landscape scenes were successively shown to Ss and syllable preferences were expressed in a paired comparison situation. Then a second series of sixty different slides of the same content proportions (Set 2 syllables, new ads and landscapes) as series 1 was shown and Ss expressed their preferences by means of a ranking procedure. The position of the syllables relative to the ads and landscape scenes was the same for all subject groups though the starting point in the sequence was randomly varied. Each series began and ended with a landscape scene. Each slide was on the screen for approximately fifteen seconds. Immediately after exposure to the first slide series, Ss were called individually from their seats to the back of the room. Each S was told she was to receive a pair of nylon stockings and that she was to choose between two brands. Two stocking boxes were shown with one of the four nonsense syllables written on each box. The S was requested to indicate the "brand" she preferred. After her choice, two more stocking boxes labeled with the remaining two syllables were presented and S was asked which of these she would choose. The four nonsense syllables were paired in all possible ways, Ss randomly assigned to comparisons. The S was

given a pair of stockings and was sent into an adjoining hall where she could not influence untested Ss.

After all Ss had chosen and received their stockings, the nonsense syllables of Set 2 were presented along with four new dummy ads and sixteen new landscape scenes. The stimulus presentation procedure was the same as for the first slide series. However, the dependent variable was not a choice between pairs of syllables but rather a rank ordering of preferences. The Ss ranked their syllable preferences from one to four on a sheet of paper passed out to each S before the experiment began.

TABLE 1
EXPERIMENTAL DESIGN

(Each cell entry is the number of times a syllable was shown to a given subject group)

| Nonsense syllables | | Subject group | | | |
|---------------------------------|--------------------|---------------|----|-----|----|
| Set 1 (paired comparison) | Set 2 (ranking) | I | II | III | IV |
| B I K | T O V | 10 | 4 | 7 | 1 |
| L U F | C I V | 1 | 7 | 4 | 10 |
| N I M | F I Y | 7 | 10 | 1 | 4 |
| G A H | R E Q | 4 | 1 | 10 | 7 |

C. RESULTS AND DISCUSSION

The choice of the syllables of Set 1 and the rankings of the syllables of Set 2 were subjected to separate analyses.

Table 2 shows the number of choices on trials 1 and 2 for the syllables of Set 1 as a function of presentation frequency. It is clear from the tabled data that the number of stimulus choices increases with increasing frequency of stimulus exposure. Chi-square analyses were conducted to test for differences between syllable frequencies. For choices on trial 1 a chi square of 7.1 ($p < .10$, 3 *df*) was obtained while for choices on trial 2, chi square was 8.2 ($p < .05$, 3 *df*). Data from trials 1 and 2 were combined by use of Pearson's $P\lambda$ Distribution (12). Differences between syllable frequencies also were significant for the combined data; chi square = 11.7 ($p < .02$, 4 *df*). No significant differences were found in number of choices as a function of the four nonsense syllables even though their association values differed (3).

Table 3 shows the effect of frequency on preference for the syllables of Set 2 when preference is indicated by rankings. These data agree with the findings of the paired comparison test. The greater the number of times a syllable was presented the higher the rank (preference) assigned to that

TABLE 2
PREFERENCE CHOICES OF NONSENSE SYLLABLES AS A FUNCTION
OF PRESENTATION FREQUENCY
(Each of 127 Ss made two choices in a paired comparison situation)

| Trial (choice) | Frequency of syllable presentation | | | | χ^2 | < | P |
|-------------------|---------------------------------------|------|-------|-----|----------|---|-----|
| | One | Four | Seven | Ten | | | |
| 1 | 22 | 35 | 28 | 42 | 7.1 | < | .10 |
| 2 | 19 | 31 | 37 | 40 | 8.2 | < | .05 |
| Combined totals | 41 | 66 | 65 | 82 | 11.7 | < | .02 |

syllable. A chi-square analysis was run for each rank separately as well as between the average rank scores of each frequency. Both analyses were highly significant and have been summarized in Table 3. Significant differences also were found between preference for the individual syllables of Set 2 (chi square = 12.0, $p < .01$, 3 *df*). This result is due largely to a preference for two of the 67 per cent association value syllables, C I V and R E Q.

TABLE 3
PREFERENCE RANKING OF NONSENSE SYLLABLES AS A FUNCTION
OF PRESENTATION FREQUENCY

| Preference rank | Frequency of syllable presentation | | | | χ^2 | < | P |
|--------------------|---------------------------------------|------|-------|-----|----------|---|------|
| | One | Four | Seven | Ten | | | |
| 1 | 12 | 23 | 38 | 54 | 32 | < | .001 |
| 2 | 11 | 36 | 42 | 38 | 19 | < | .001 |
| 3 | 34 | 45 | 27 | 21 | 10 | < | .02 |
| 4 | 70 | 23 | 20 | 14 | 63 | < | .001 |

Note: Average χ^2 = 31, $P < .001$.
N of Ss = 127.
N of observations = 508.

It is clearly evident that Ss showed a marked tendency to prefer the nonsense syllables they had seen more frequently. Although we do not contend that word "value" as opposed to word "goodness" was measured in the present study, our results do indicate that the effect of exposure frequency upon preference for nonsense syllables is not limited to measures obtained with the semantic differential. Thus the present experiment extends and lends support to the previous study by Johnson, Thomson and Frincke (7).

Postman (11) in reviewing the literature concerned with the relation between word value and word frequency has stated, "Finally, we cannot entirely discount the possibility that familiarity resulting from frequency

may be in itself a source of positive value" (p. 68). The data from the present experiment, along with results of earlier studies (5, 7, 9, 10), lead us to conclude that stimulus familiarity is indeed an important determiner of stimulus preference.

D. SUMMARY

It was predicted that Ss would prefer nonsense syllables they had seen frequently. Positive results were obtained.

REFERENCES

1. ADAMS, H. F. Memory as affected by isolation of material and by repetition. *J. Appl. Psychol.*, 1927, **11**, 25-52.
2. ARCHER, E. J. Re-evaluation of the meaningfulness of all possible CVC Trigrams. *Psychol. Monog.*, 1960, **74** No. 10 (Whole No. 497).
3. GLAZE, J. A. The association value of nonsense syllables. *J. Genet. Psychol.*, 1928, **35**, 255-269.
4. GOLDBLUM, I. Indicators of perception: I. Subliminal perception, subception, unconscious perception: An analysis in terms of psychophysical indicator methodology. *Psychol. Bull.*, 1958, **55**, 373-411.
5. HARTLEY, E. L. The learning of tastes. *Publ. Opin. Quart.*, 1960, **24**, 621-631.
6. JERSILD, A. Primacy, recency, frequency, and vividness. *J. Exper. Psychol.*, 1929, **12**, 58-70.
7. JOHNSON, R. C., THOMSON, C. W., & FRINCKE, G. Word value, word frequency and visual duration thresholds. *Psychol. Rev.*, 1960, **67**, 332-342.
8. ———. A response to Stricker's comments on "word values, word frequency, and visual duration thresholds." *Psychol. Rev.*, 1962, **69**, 239-240.
9. KRUGMAN, H. E. Affective response to music as a function of familiarity. *J. Abn. & Soc. Psychol.*, 1943, **38**, 338-392.
10. MASLOW, A. H. The influence of familiarization on preference. *J. Exper. Psychol.*, 1937, **21**, 162-180.
11. POSTMAN, L. The experimental analyses of motivational factors in perception. In J. S. Brown (Ed.), *Current Theory and Research in Motivation*. Lincoln: Univ. Nebraska Press, 1953. Pp. 59-108.
12. RAO, C. R. Advanced Statistical Methods in Biometric Research. New York: Wiley, 1952.
13. STRICKER, G. Word values, word frequency, and visual duration thresholds: A comment. *Psychol. Rev.*, 1961, **68**, 420-422.
14. STRONG, E. K. The effect of size of advertisements and frequency of their presentation. *Psychol. Rev.*, 1914, **21**, 136-152.
15. WILSON, W. R., & BECKNELL, J. C. The relation between the association value, pronunciability, and affectivity of nonsense syllables. *J. of Psychol.*, 1961, **52**, 47-49.

Advertising Department

E. I. du Pont de Nemours & Company
Wilmington 98, Delaware

THE PRODUCTION OF RANDOM LETTERS SEQUENCES IN SCHIZOPHRENICS*

Norwich Hospital, Norwich, Connecticut

HALIDE S. YAVUZ

A. INTRODUCTION

Baddeley (2) recently reported the results of several experiments in which subjects undertook to produce letters of the alphabet in a random order. The subjects were instructed to imagine they were drawing letters one by one from a hat, calling them, and replacing them. They showed a strong tendency to repeat pairs of letters, especially those comprising adjacent letters in the alphabet such as AB, BC, XY, etc. The tendency toward repetition was increased when the rate of responding was speeded and when the subjects performed a distracting task. This paper reports the results of a pilot study involving the use of Baddeley's method with schizophrenics and normal subjects. It was hypothesized that schizophrenics, when given the Baddeley task, should exceed normal subjects in the tendency to repeat letter pairs. It is evident that bigrams comprising adjacent letters of the alphabet and letters sequences commonly occurring in words must represent old and well-established habits. The degree of success of subjects in emitting letters in a random order must depend appreciably on their avoidance of the old habits. Reviews of the theories of schizophrenias such as undertaken by Arieti (1) indicate general support of the assumption that anxiety and consequent regressive behavior are characteristic of this psychosis. Such theorizing supports the deduction that the schizophrenic because of his anxiety potential should show a strong tendency, when instructed to perform the Baddeley task, to regress to his old and well-established letter sequence habits.

B. METHOD

The subjects comprised 13 hospitalized schizophrenics in a state of remission and a comparison group of 12 psychiatric aides. The subjects of the two groups were equated insofar as possible on the basis of age, education, and clinically estimated intelligence. Each subject was instructed to call

* Received in the Editorial Office on April 17, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

out letters of the alphabet in a random or completely mixed up way. They were further told to say one letter every second in time with the ticking of a metronome. In this way each subject produced three sequences of 100 letters per run with a rest of two minutes between runs.

C. RESULTS AND DISCUSSION

The 300 letters emitted by each subject were tabulated on a 26×26 matrix representing all the letters of the alphabet. This matrix indicated all 676 of the possible sequences of letters in pairs. The tabulations were such as to provide the basis for computing a variety of scores, two of which are reported here. The first was the number of repetitions of letter pairs or bigrams for each subject. In these terms when the same bigram occurred twice it was counted as one repetition; if it occurred three times, it was counted as two repetitions, etc. The number of repetitions counted in this way for the two combined groups of subjects ranged from 39 to 148 with a median of 66. Among the schizophrenics, eight were above and four below the median, whereas, for the normal subjects, three were above and nine below. Siegel's (3, p. 111f) median test shows a significant difference in the predicted direction of more repetitions for the schizophrenics than for the normal subjects, $p < .05$. With the Yates correction, however, the p falls between .25 and .10. Fisher's exact test (3, p. 96f) yields a p of .0511.

There was also a tendency on the part of the schizophrenics to produce more repetitions of sequences of adjacent letters of the alphabet than was found in the normal subjects. This tendency, however, fell short of statistical significance. The range of the number of these sequences for combined groups was from 6 to 116 with 21 as the median. Seven of the schizophrenics were above and five were below this median; the normal group had five above and five below.

The results of this exploratory study, while not statistically conclusive, indicate a proneness on the part of schizophrenics to regress to the use of old habits when they are instructed to perform a simple task calling for the avoidance of these habits. The Baddeley method is a new device which may be used for studying these tendencies. It has the advantage of simplicity and it provides objectively scorable data. Several comments may be made regarding possible refinements in the method as it was employed in the present study. The schizophrenic subjects were more prone than the normals to block and delay during their emission of the letter. Without such blocking and delaying it should be possible to produce 100 letters in one minute and 40 seconds. When it became evident that the schizophrenics were

showing a strong proneness to falter in the task, timing of the performance was undertaken. One of the patients required eleven minutes for a single run of 100 letters. It is evident that the task should be timed and the interruptions in the production of letters should be recorded. Furthermore, these factors should be taken into account in the treatment of the data. As already implied, the results obtained by the Baddeley method can be analyzed in many different ways. These merit exploration. The problem of the choice of subjects merits comment. In retrospect, it appears that better precautions than were taken in the present study are needed in order to insure homogeneity in schizophrenic subjects. These problems appear capable of resolution and the findings reported justify further research with the Baddeley method.

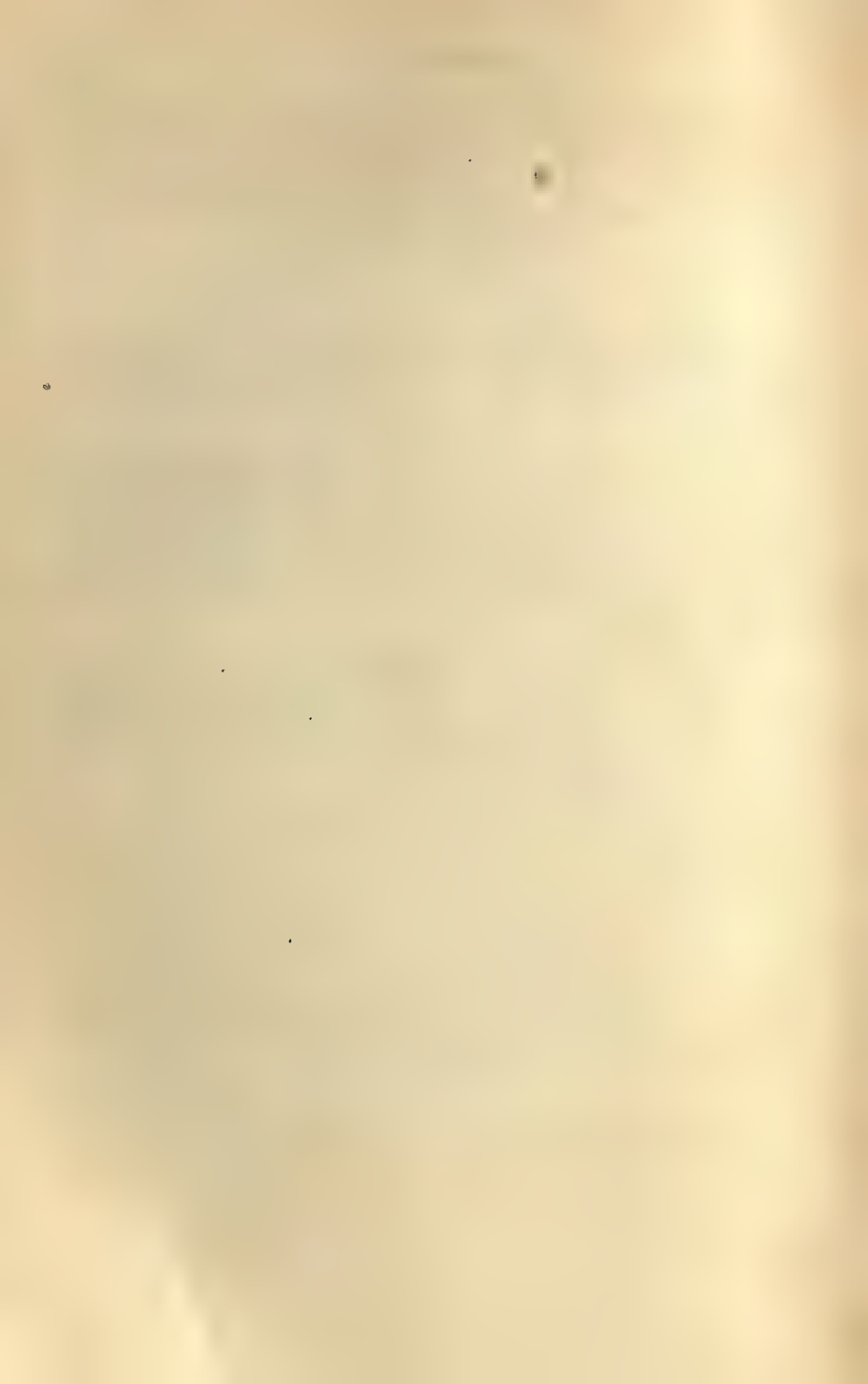
D. SUMMARY

This was an exploratory study making use of Baddeley's method of having subjects attempt to produce letters of the alphabet in a random order. Each of 13 schizophrenics and 12 hospital aides produced three 100-letter runs of letters. The data indicated a tendency for the schizophrenics to exceed the aides in repeating pairs of letters, especially pairs comprising adjacent letters in the alphabet.

REFERENCES

1. ARIETI, S. Interpretation of schizophrenia. New York: Brunner, 1955.
2. BADDELEY, A. D. Some factors influencing the generation of random letter sequences. Paper presented at the London, England, meeting of the Experimental Psychology Society, January, 1962.
3. SIEGEL, S. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill, 1956.

*Psychological Laboratories
Norwich Hospital
P.O. Box 508
Norwich, Connecticut*



BRIGHTNESS FROM REPETITIVE PHOTIC TRAINS VARYING
IN PULSE NUMBER AND SEPARATION INTERVALS:
SENSORY IMPLICATIONS OF THE ALTERNATION
OF RESPONSE THEORY, III*¹

Department of Psychology, Michigan State University

THOMAS M. NELSON, S. HOWARD BARTLEY, AND ROBERT M. JEWELL

A. PROBLEM

Brightness can be varied by manipulating time characteristics of stimulation, intensity constant. There are three general sets of experimental circumstances in which manipulation of timing has been found effective. The first is associated with observation of prolonged intermittent photic impingements, the second with observation restricted to several photic repetitions, and the third with observation of single pulses.

The first set includes phenomena following the Talbot-Plateau law, Bartley's alpha effect, and a more general range of sensory outcomes called the Brücke and Bartley or brightness enhancement effects. Enhancement means brightness greater than that produced by extended steady illumination.

The following general facts are known. One, if the composition of the cycle [the pulse-to-cycle fraction (PCF)]² is held constant, and rate of photic intermittency increased until flicker disappears, the brightness of the target will be equal to that of a constantly illuminated target of the mean intensity of the intermittent. This is the Talbot-Plateau law (see Fig. 1A). Two, when rate of photic intermittency is adjusted to approximately intrinsic cortical rhythm, brightness will vary somewhat inversely with the PCF (Bartley's alpha effect, Fig. 1B). Between 30 and 150 c/ft², for example, small PCFs in the range of .20 to .35 will produce the greatest brightnesses as evidenced by the greatest ratios of steady to intermittent illumination to produce a match. Lesser and greater PCFs reduce the effectiveness of stimulation. However in the range of intensity below this,

* Received in the Editorial Office on April 18, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ The present paper is a part of the work being done under Research Grant NSF-G-19485 from the National Science Foundation.

² Pulse-to-cycle fraction is designed to replace the "light-dark ratio," which is a response description of cycling. The older terminology did not distinguish between sensory outcome and sensory input and hence led to misunderstandings.

maximal effect is produced by a larger PCF (3). Three, intermittency rates below fusion can produce brightnesses greater than that of continuous illumination of the same intensity (Brücke and Bartley effects of brightness enhancement). When PCF is fixed and intensity reduced but above a critical value, enhancement will occur with the slower rates (Fig. 1C) (1).

In the *second* set of general circumstances, brightness is varied by varying the number of cycles of repetition of stimulation in a single train when the PCF and cycle length are fixed. Little has been done along this line, but because short trains of pulses produce CFF with greater pulse separations than longer trains (Fig. 1D) (2), one can infer that intermittency rates giving fusion for the shorter trains will be less bright than longer trains made intermittent at the same rate (Fig. 1E). Also the results of one investigation show that short pulse trains will produce greater brightness enhancement than long pulse trains when stimulation is cycled at 10 per sec., *i.e.*, at approximately alpha rate (Fig. 1F) (6).

There are two phenomena associated with the *third* set of circumstances, observation of single pulses. One, the classical Bunsen-Roscoe law states the nature of the relation between the time value of the photic stimulation and its brightness as a reciprocity. For pulses of moderate intensity and very short duration, the product of intensity and duration is a constant. Two, the brightness of longer photic pulses is related to the null period preceding the pulse. The precise quantitative character of the changes is not known nor has a wide range of conditions producing increases been studied. However, with pulse duration fixed in the range of .33 to .67 msec and intensity in the range of 3.6 c/ft² to 36.3 c/ft², effects of enhancement magnitude will be produced by separating successive pulses by 500 msec or more (5). At somewhere near 2000 msec, further lengthening of the null interval produces no further increase of brightness. It is at this point that one ceases to deal with effects of one pulse upon another and thus begins to deal with truly *isolated pulse* effects (5).

Brightness effects evoked by temporal manipulation of radiance depend upon the neurophysiological capabilities of the optic pathway. Because brightness depends upon the activity of separate but related mechanisms, it is possible that brightness evoked by one means will integrate with that evoked by another. If effects from varying the null interval are integrated with those from varying the number of cycles of stimulation, the brightness of pulse trains of given lengths will increase as the separation interval increases (5) while the effect of the separation interval upon brightness will decrease as the number of cycles in the train increases (6). The investigation to be

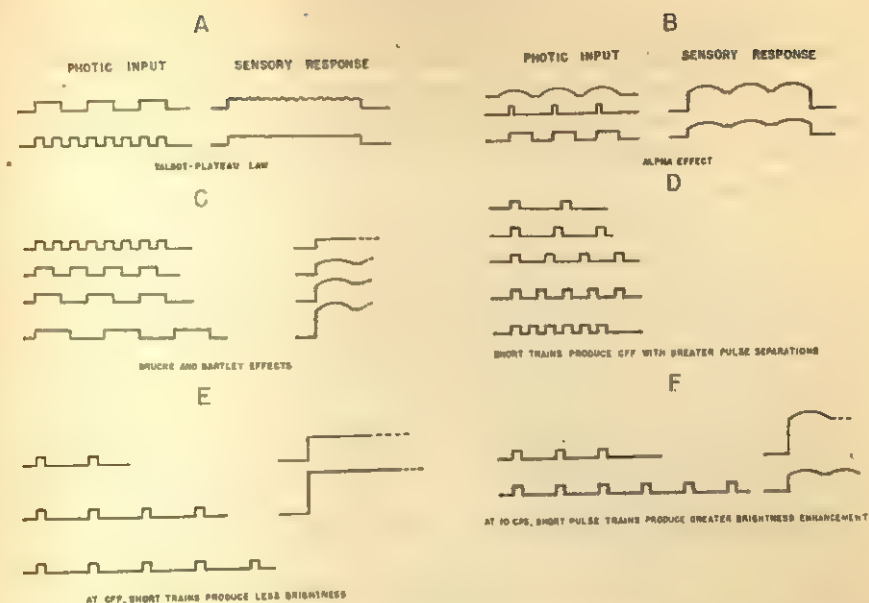


FIGURE 1

DIAGRAMMATIC REPRESENTATION INTERRELATING PHOTIC INPUT AND SENSORY RESPONSE

A. Input rate can be made high enough to produce steady brightness. B. Increasing pulse length (raising PCF) decreases brightness when rate is held constant and in time with the cortical rhythm (alpha waves). C. As input rate is reduced from CFF, brightness increases, although some flicker is introduced. D. A series of trains from 2 to 5 pulses all producing CFF. Pulse separations for short trains must be greater than for long. E. Short trains at CFF produce lower brightness than long. F. At 10 cps reverse is true.

reported bears upon this possibility. In this experiment short pulse trains were separated by intervals of various lengths and their brightness matched with that of a steady target.

B. METHOD

The episotister, photic sources and pulse form used in this study have been described elsewhere (5, 6). The experimental and matching targets used were rectangular and did not fall wholly on the fovea. At observation distance each subtended $2^{\circ} 5'$ in the horizontal direction, $45'$ in the vertical, and were separated in the horizontal direction by $36'$. Levels of I used were 36.3 c/ft^2 and 3.6 c/ft^2 . The procedure and Os were as in preceding studies (5, 6).

Photic pulses were grouped in 3-, 5- and 9-unit trains. Pulse trains

were cycled so as to give a PCF of either .33 or .66, *i.e.*, pulses 33 msec long followed by null periods of 66 msec, and pulses 66 msec long followed by a null period of 33 msec, respectively. However, the cycle lengths *within* every train were the same regardless of number of cycles. There was no modulation within the separate trains. Stimulus rate within the trains was fixed at 10 cps.

The following intervals *between* trains of pulses were used. Five interval conditions were used for 3-unit trains when PCF was .33. Those used were 166 msec, 366 msec, 566 msec, 966 msec and 2166 msec, respectively. Five-unit trains were separated by intervals of 366, 766, and 1966 msec respectively; and 9-unit trains by 366 and 1566 msec, respectively when PCF was .33. Fig. 2 depicts the PCF .33 conditions and is included in order to clarify the nature of the experimental operations. When PCF was .66, three-unit trains were separated by 133, 333, 533, 933 and 2133 msec, respectively; 5-unit trains by 333, 733, and 1933 msec, respectively; and 9-unit trains by 333 and 1933 msec respectively.

C. EVALUATION OF SENSORY RESPONSE

The data are described in terms of a brightness index which is a ratio of the mean intensity of brightness matchings for the brief pulses to the mean intensity of matchings of the same target when continuously illuminated. An index of 100 represents a match when the intensities of the steady and intermittent presentations are equal; 50, a match when the steady target had to be half as intense as the intermittent; 200, twice as intense, etc. All points above 100 therefore indicate the occurrence of brightness enhancement or the *Brücke* and *Bartley effects*, *i.e.*, that the brief photic pulse is brighter than steady illumination of the same intensity. It should be pointed out in passing, however, that the index of brightness was devised simply to allow comparisons of effects induced by dissimilar conditions. The indexes are ordinal measurement and not additive in the strict sense. An index of 200 therefore does not necessarily mean that the briefly exposed target looked twice as bright as the steady, but rather that the intensity of the steady target had to be doubled for the two to match. Likewise indexes of 200 do not indicate the same relative increase of brightness when the ranges of intensity matched differ.³

³ In Fechnerian psychophysics, sensation changes and intensity changes have a logarithmic relation. In Stevens' scheme a power function governs the relation. In no case do the changes bear a simple linear relation to one another.



FIGURE 2

DIAGRAMMATIC REPRESENTATION OF THOSE CONDITIONS OF PHOTIC INTERMITTENCY ASSOCIATED WITH PCF .33
 Each dash represents a pulse duration of 33 msec, and the shortest interspaces a null interval of 66 msec. The intervals separating the trains are represented as longer spaces and their value given in the representation itself. The PCF .66 condition is described in the text.

TABLE 1

BRIGHTNESS INDEXES FOR DIFFERENT NUMBERS OF PHOTIC REPETITIONS, PCFs, AND SEPARATING INTERVALS

Indexes are stated as percentage above or below the index value of 366 msec when PCF is .33, and 333 msec when PCF is .66.

| PCF | Null period in msec | Low I (3.6 c/ft ²) | | | | | High I (36.3 c/ft ²) | | | | |
|----------|---------------------------|-----------------------------------|-----|------|------|------|-------------------------------------|------|------|------|------|
| | | 166 | 366 | 566 | 966 | 2166 | 166 | 366 | 566 | 966 | 2166 |
| Observer | | | | | | | | | | | |
| 3 pulse | N | -8.5 | 000 | 6.4 | 19.7 | 37.3 | -31.9 | 000 | 0.8 | 20.2 | 2.6 |
| | BJ | 9.7 | 000 | 7.2 | 31.1 | 57.8 | 54.3 | 000 | 12.0 | 4.7 | 34.0 |
| | | | 366 | 766 | 1966 | 366 | 766 | 1966 | | | |
| 5 pulse | N | | 000 | 15.3 | 37.4 | | | 000 | 50.4 | | 37.8 |
| | BJ | | 000 | 12.4 | 54.0 | | | 000 | 15.0 | | 18.3 |
| | | | 366 | 1566 | 366 | 1566 | | | | | |
| 9 pulse | N | | 000 | | 17.3 | | | 000 | | | 20.8 |
| | BJ | | 000 | | 43.3 | | | 000 | | | 0.7 |
| | | | | | | | | | | | |
| PCF | Null period in msec | 133 | 333 | 533 | 933 | 2133 | 133 | 333 | 533 | 933 | 2133 |
| Observer | | | | | | | | | | | |
| 3 pulse | N | -20.0 | 000 | 30.8 | 64.3 | 76.5 | -10.4 | 000 | 18.9 | 37.7 | 28.3 |
| | BJ | 3.1 | 000 | 14.5 | 55.6 | 87.2 | - 3.0 | 000 | 13.8 | 15.7 | 14.1 |
| | | | 333 | 733 | 1933 | 333 | 733 | 1933 | | | |
| 5 pulse | N | | 000 | 20.2 | 33.0 | | | 000 | 13.5 | | 14.9 |
| | BJ | | 000 | 17.8 | 39.6 | | | 000 | 30.7 | | 13.1 |
| | | | 333 | 1533 | 333 | 1533 | | | | | |
| 9 pulse | N | | 000 | | 13.6 | | | 000 | | | 10.4 |
| | BJ | | 000 | | 37.5 | | | 000 | | | 12.9 |
| | | | | | | | | | | | |

D. RESULTS AND DISCUSSION

The outcomes for the 36.3 c/ft² conditions are presented graphically in Figure 3 and the 3.6 c/ft² conditions in Figure 4. The data upon which the graphs are based are in Table 1 except for the curves depicting the 1-pulse conditions. These were computed from data presented in an earlier paper. All curves have been adjusted to coincide at one of two common index values. Where PCF is .33 the value is that of the 3-pulse train at a separation value of 366 msec. Where PCF is .66 it is that of a 3-pulse train having a separation value of 333 msec. Pegging the curves at a common value enables ready comparison of present and previous results (5) with respect to form and rate of change, while not altering relationships within the data.

As can be seen, the data support the hypotheses stated in the last paragraph

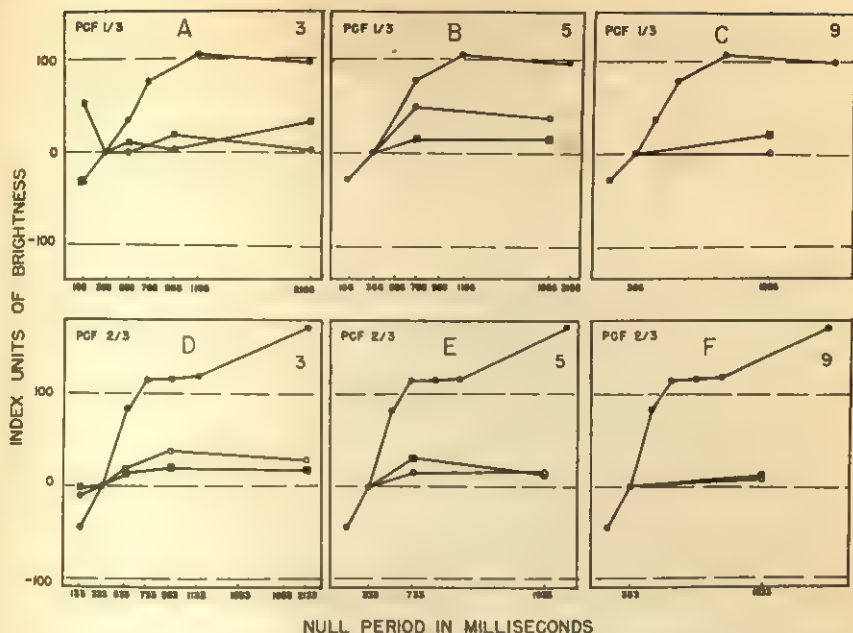


FIGURE 3

BRIGHTNESS INDEXES ASSOCIATED WITH INCREASES IN NULL INTERVAL TIME

All indexes were plotted so as to converge at a common null value of 366 msec for a PCF of .33 and 333 msec for PCF .66. The open circles represent data from the principal O and the closed squares from the student Os. Numbers appearing in the upper right corner indicate whether the preceding points are for 3, 5, or 9 photic repetitions. The closed circles are based upon brightness matchings of single pulses (5). Intensity was fixed at 36.3 c/ft² for all curves.

of section A. First, there is an increase in brightness of the 3-, 5- and 9-unit trains as separation time is increased. Changes in value are in the same direction as for the 1-pulse condition even though trains of pulses are being employed rather than single pulse units. Secondly, in every case the separation between trains produces a less steep rise in brightness when pulse trains are used than is the case where single pulse units are employed.

Not explicitly predicted, but nevertheless also evidenced from the fact that effects from one mechanism serving brightness will integrate with another, is that variation in the interval separating trains of pulses varies brightness less than does the same time variation between single pulses. Notice that curves of each graph of Figures 3 and 4 are equated at either the 366 or 333 msec separations but that the curves for the 1-pulse condition attain the highest value. This structuring presumably occurs because the

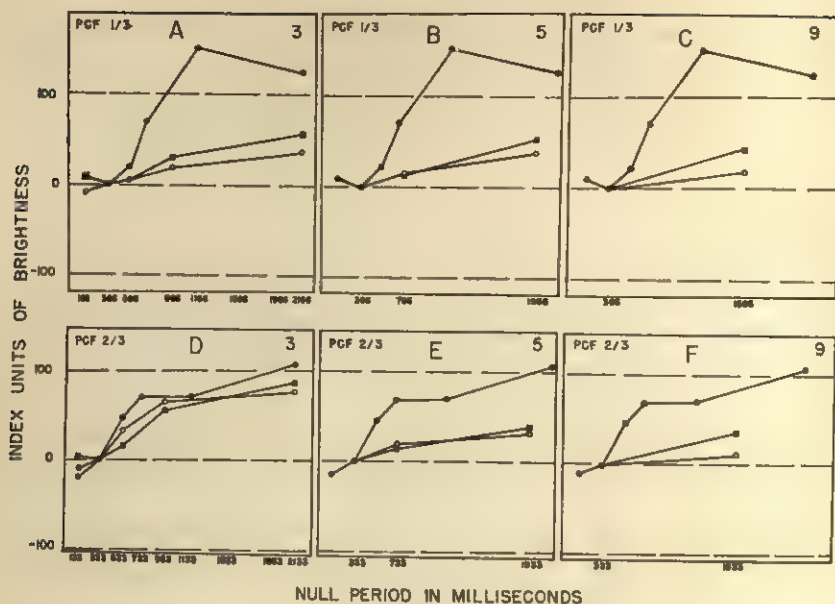


FIGURE 4

BRIGHTNESS INDEXES ASSOCIATED WITH INCREASES IN NULL INTERVAL TIME

All indexes were plotted so as to converge at a common null value of 366 msec for a PCF of .33 and 333 msec for PCF .66. The open circles represent data from the principal *O* and the closed squares from the student *O*s. Numbers appearing in the upper right corner indicate whether the preceding points are for 3, 5, or 9 photic repetitions. The closed circles are based upon brightness matchings of single pulses (5). Intensity was fixed at 3.6 c/ft² for all curves.

effect of the separation period is largely upon the nervous discharge to the first pulse of the train. The effect upon the initial spike is increased as temporal separation between trains is increased, but since brightness of the train is associated with the average level of nervous activity during the period of stimulation and this level of activity decreases past the first pulse, brightness will never attain the magnitude produced by single pulses having the same temporal separation.

Finally, it might be mentioned that in dealing with pulse trains in which the separation between trains affects the brightness, we are in fact dealing with stimulus *modulation* and not simply stimulus repetition involving a fixed number of cycles. We are doing so just as certainly as we are dealing with effects from repetitive and not isolated stimulation where the length of the null interval between pulses affects the brightness of the pulse (5). However, discussion of brightness effects in modulation terms demands the

presence of a more comprehensive means of describing repetitive stimulation than we now use. Such a descriptive device not only needs to deal with complex temporal characteristics of radiance, but must also relate to known or *likely* biological events, *viz*, neurophysiological processes. It is therefore necessary that the device be natural (4) and not achieved by imposing a generalized mathematical model upon the phenomena, *e.g.*, analysis by Fourier procedure. The description should itself give as many hints as possible as to how the visual mechanism actually attains known sensory end results. A device fulfilling these requirements does not seem extant but is now in the preliminary stage of development.

REFERENCES

1. BARTLEY, S. H. Brightness enhancement in relation to target intensity. *J. of Psychol.*, 1951, **32**, 57-62.
2. BARTLEY, S. H., NELSON, T. M., & RANNEY, J. E. The sensory parallel of the reorganization period in the cortical response in intermittent retinal stimulation. *J. of Psychol.*, 1961, **52**, 137-147.
3. BARTLEY, S. H., PACZEWITZ, G., & VALSI, E. Brightness enhancement and the stimulus cycle. *J. of Psychol.*, 1957, **43**, 187-192.
4. NELSON, T. M., & BARTLEY, S. H. Numerosity, number, arithmetization, measurement, and psychology. *Philos. Sci.*, 1961, **28**, 178-203.
5. NELSON, T. M., BARTLEY, S. H., & JEWELL, R. M. Effects upon brightness produced by varying the length of the null interval separating successive "single" pulses: Sensory implications of the alternation of response theory, I. *J. of Psychol.*, 1963, **56**, 99-106.
6. ———. Brightness changes associated with variation in the number of pulses in a repetitive photic train: Sensory implications of the alternation of response theory, II. *J. of Psychol.*, 1963, **56**, 107-113.

Department of Psychology
Michigan State University
East Lansing, Michigan



HSPQ NORMS ON A STATE-WIDE DELINQUENT POPULATION*

*Green Hill School, Chehalis, Washington and Fort Worden School, Port Townsend,
Washington*¹

GEORGE R. PIERSON AND ROBERT F. KELLY

A. INTRODUCTION

There has been considerable recent interest in the application of factored personality tests to the study of delinquency and related problems of adolescence (see e.g., 2, 5, 7). The Handbook of the High School Personality Questionnaire (2) by IPAT² (subsequently referred to as the HSPQ) presents a delinquent personality profile (see Table 1) showing teenage delinquents to be below average on Factor A, or cold and aloof; below average on Factor B, or less intelligent; less dependable and conscientious as measured by below average scores on Factor G; more dominant and aggressive by above average scores on Factor E; and more obstructively individualistic by an above average Factor J score (2, p. 21).

Preliminary analysis of over 1300 HSPQ profiles of delinquents committed to the State of Washington's Bureau of Juvenile Rehabilitation has begun, and means, standard deviations, and intercorrelations of the 14 factors have been calculated.³ While these findings are in some respects consistent with those quoted above, there are certain essential differences which merit report. This sample, while representative of the delinquent population of the State of Washington is not, perhaps, representative of the megapolitan areas. It does contain representatives of the large city, however, and would appear representative of a very large segment of the American delinquent population. This note is offered for those working with the problem of delinquency and its diagnosis.

* Received in the Editorial Office on April 19, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ Green Hill School, Chehalis, Washington and Fort Worden School, Port Townsend, Washington are institutions of the Bureau of Juvenile Rehabilitation, Department of Institutions, Washington.

² Institute for Personality and Ability Testing, Champaign, Illinois.

³ The data for this study were gathered at the Fort Worden Diagnostic Center in connection with another study, and were processed at the Pacific Northwest Computer Laboratory at the University of Washington.

B. METHOD

The HSPQ was administered routinely to all male juveniles, aged 14 years and over, received at the Fort Worden Diagnostic and Treatment Center. Utilized in this study were profiles of all boys received from the courts between April, 1961, and October, 1962. The administration was done in groups of eight to 15 subjects within the diagnostic residence cottages. Time of administration was the fourth week of residence in a six-week stay.⁴ Initially only Form A was administered due to time limitations, with Form B being included on those boys received after April, 1962. This resulted in 531 boys with Form A only and 339 with both forms.

Those boys unable to read at or above the 5th grade level as measured by the Stanford Achievement Test, Form J, were given the HSPQ orally in separate groups. (Cattell by personal communication raised the question of the effects of "violation of individuality of answers" here. These *Ss* are nevertheless included in the total because the investigators estimated no serious distortion, if such effects were present, would result since small sample checking disclosed no mean difference between types of administration.)

The raw scores were calculated and transferred to IBM cards, one card for each subject on each form of the HSPQ. A program was written for, and calculations performed by, an IBM 709 data processing system. The means, *SD* and intercorrelations were obtained. The means and *SDs* are presented and discussed below.

C. RESULTS AND DISCUSSION

Table 2 shows the means and standard deviations of this sample for forms A and B, together with the means and standard deviations of the reported standardization sample. It will be noted that no differences as great as those quoted in Table 1 were found.⁵ An explanation for this will be offered later, but first a number of discrepancies between the present findings and those in the HSPQ Handbook seem to have considerable theoretical significance and should be discussed.

Significantly higher scores on Factors A and H have now been found which suggest a more outgoing, adventurous, extraverted personality than that indicated by Table 1. However, in addition to the adjectives *warm* and

⁴ This was an arbitrary decision for administrative reasons, and since has had to be changed to the second week of residence.

⁵ Although Table 1 is in *Stens* and Table 2 in raw scores, comparisons are made throughout in terms of significant differences above or below mean standardization scores—the direction of which is clinically important.

TABLE 1
PERSONALITY FACTOR PROFILE OF DELINQUENTS
(In Stens)

| A | B | C | D | E | F | G | H | I | J | O | Q2 | Q3 | Q4 |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 5.2 | 4.0 | 5.1 | 6.0 | 6.4 | 5.3 | 4.7 | 5.4 | 5.3 | 6.5 | 5.4 | 5.8 | 4.8 | 5.8 |

Note: $N = 413$. In distribution of Stens, $M = 5.5$, $SD = 2.0$.

sociable used to describe the high A individual, the following are also applicable: easy-going, casual, careless, and readily laughs. Also, high scores on Factor H describing the socially desirable quality of liking to meet people, also mean adventurous, impulsive, frivolous, careless, and especially does not see danger signals. These descriptive terms would in themselves seem to describe delinquent behavior in general.

Further, the long-held view that delinquents are below average in intelligence cannot be accepted, particularly the kind of intelligence measured by the HSPQ Factor B, which is relatively free of cultural and educational contamination. Admittedly the delinquents of our population have been found to be extremely poor readers. Earlier studies have shown this problem to be attributed in part to high Factor H, which is most significantly above average in our population (4). In a sample of 127 delinquent boys the mean paragraph meaning grade placement, as measured by the S.R.A. Diagnostic Reading Test, was seven years, six months, showing them to have an average of two years, three months reading deficit.⁶ With respect to intelligence, however, a random sample of intelligence quotients measured by either the WISC or the Wechsler Adult Intelligence Scale and calculated at Green Hill School showed a mean *IQ* of 98.5.⁷ This corroborates a previous sample reported by Tyler and Kelly (8). Furthermore, evidence of the intellectual normalcy of *delinquents* is found on Factor B shown in Table 2.

Despite what would seem to be theoretically sound reasoning, we do not find a defective superego to be characteristic of the present delinquent population, as herein assessed, since the Factor G mean is not significantly different from the normal population. The reported findings that criminals and psychopaths are low on Factor G (2, p. 22) would be consistent with a strongly held clinical opinion of workers at these facilities that delinquents are not for the most part criminal types or psychopaths, but rather a unique kind of pathology that we are not prepared to define at this time.

⁶ The data for this report were gathered by Mr. Gordon Hey, principal of the academic school at Green Hill School, Washington.

⁷ These measures of intelligence were done as part of the diagnostic program of the Fort Worden Diagnostic and Treatment Center, Washington.

TABLE 2
COMPARISON OF IPAT AND BJR NORMS
HSPQ Factor Raw Score Averages with Differences

| | A | B | C | D | E | F | G | H | I | J | O | Q2 | Q3 | Q4 |
|---------------------------------------|----------------|-----|------|------|------|------|------|------|------|------|-----|------|------|------|
| Form A (<i>N</i> = 870) ^a | | | | | | | | | | | | | | |
| BJR | Score 5.3 | 6.8 | 5.8 | 4.7 | 4.9 | 5.0 | 6.2 | 5.4 | 3.9 | 5.7 | 4.8 | 5.5 | 4.4 | 3.9 |
| | <i>SD</i> 1.9 | 1.9 | 1.8 | 1.8 | 1.9 | 1.7 | 1.8 | 2.1 | 1.8 | 1.5 | 1.8 | 1.5 | 1.6 | 1.7 |
| IPAT | Score 4.9 | 6.7 | 5.4 | 5.0 | 5.0 | 5.4 | 6.0 | 4.7 | 4.2 | 5.1 | 5.0 | 5.5 | 4.9 | 4.5 |
| | <i>SD</i> 1.7 | 1.9 | 1.9 | 1.7 | 1.9 | 1.7 | 1.8 | 2.1 | 2.1 | 1.6 | 1.8 | 1.8 | 1.7 | 1.7 |
| Diff. | Score +3** | +1 | +4** | -3** | -1 | -4** | +2 | +7** | -3** | +6** | -2 | .0 | -5** | -6** |
| Diff. | <i>SD</i> +2** | .0 | -1 | +1 | .0 | .0 | .0 | .0 | -3** | -1 | .0 | -3** | -1 | .0 |
| Form B (<i>N</i> = 339) ^b | | | | | | | | | | | | | | |
| BJR | Score 6.3 | — | 5.2 | 3.5 | 4.1 | 4.3 | 6.1 | 4.8 | 5.6 | 4.5 | 4.9 | 5.1 | 5.6 | 4.0 |
| | <i>SD</i> 1.9 | — | 1.8 | 1.7 | 1.7 | 1.8 | 2.1 | 1.9 | 1.7 | 1.6 | 1.9 | 1.5 | 1.6 | 1.7 |
| IPAT | Score 5.4 | — | 5.1 | 4.4 | 4.8 | 4.5 | 5.8 | 4.8 | 5.1 | 4.4 | 5.0 | 5.4 | 5.8 | 4.2 |
| | <i>SD</i> 1.9 | — | 1.8 | 1.6 | 1.8 | 1.7 | 2.0 | 2.0 | 1.6 | 1.6 | 1.8 | 1.6 | 1.6 | 1.7 |
| Diff. | Score +.9** | — | +1 | -9** | -7** | -2 | +3** | .0 | +5** | +1 | -1 | -3** | -2 | -2 |
| Diff. | <i>SD</i> .0 | — | .0 | +1 | -1 | +1 | +1 | -1 | +1 | .0 | +1 | -1 | .0 | .0 |

^a On Factor B, *N* = 465.

^b Factor B omitted on last 405 *Ss* of sample.

** Significant at .01.

Close observation of delinquents suggests that they are more dominant and aggressive than average, so Cattell's reported elevation of Factor E would hardly seem surprising. Again, however, our findings do not show significantly high scores on this factor. It is possible that what is seen and interpreted as dominance is actually a combination of other factors at work, possibly Factor A and especially the exuberance found in high Factor H, combined with other factors. Also, the rebellion indicated by a high Factor J could possibly be misinterpreted as dominance.

Factor I in this sample presents equivocal evidence that is difficult to interpret. In the earlier work with delinquents, as examination of Table 1 will show, nothing unusual was found. Low Factor I is described as tough, realistic, and showing few artistic responses. Conversely, high Factor I indicates esthetic sensitivity and "tendermindedness." The low Factor I boy would seem to be a masculine individual of adequate psychosexual identification. Form A of the HSPQ is significantly below average on this factor. Yet one can hardly read a historical or psychological report of a delinquent without seeing some reference to inadequate psychosexual identification. This finding that the boys of this sample, as measured by HSPQ Factor I, Form A, are below average would seem to indicate that they are compensating for this hypothesized weak masculine identity. In fact, the entire delinquent façade or delinquent identification seems to serve the purpose of proving to the delinquent that he is really a tough man. This interpretation is increasingly credulous when it is considered that Factor I has been found to increase significantly in delinquent boys following psychotherapy (6). This was interpreted to mean that they no longer needed the hypermasculine façade and reported themselves more realistically near the population mean on this factor. This does not, unfortunately, account for the findings on Form B (Table 2) on which Factor I was found significantly above the population mean. It is again possible that this discrepancy could be accounted for on the basis of delinquent types, but only further study will clarify the matter. It will be noted that Factor I on Form B of the standardization sample was significantly higher than Form A, as is so on the present sample. Since the direction is the same, it is speculated that the greater magnitude of difference on this factor is a statistical artifact associated with the present sample.

A recent paper by Grossbard (3), has received considerable attention. Grossbard offers the theory that the delinquent has a defective ego and offers several behavioral observations to support his view. Present findings do not support Grossbard's theory. In view of the size of this sample it must be

concluded that what is rated behaviorally as weak ego is the combination of other factors that resemble the ego-strength factor. The difficulty of evaluating Grossbard's observations is that he fails to define operationally what he means by Ego. Ego strength is herein defined as that which is measured by Factor C of the HSPQ, and other IPAT instruments. Since the present sample shows significantly higher scores on this factor, it is suggested that low score on Factor Q₃ could explain the discrepancy between Grossbard's observations and the present findings. Factor Q₃, an integration factor, has been called by Stice (HSPQ Handbook) the "gyroscopic" factor. It describes at the high end the incorporation of socially approved character responses, self-control, persistence, foresight, considerateness of others, and conscientiousness; and could be a part of what is often reported clinically as ego strength.

The second order factor of anxiety is below average in this sample and it is not surprising that Factor Q₄, a component of anxiety, is also significantly low. However, a reasonable hypothesis would have been in the absence of these data, that delinquents are high on ergic tension or inner turmoil factor. However, the present population definitely do not show this at time of testing.

An explanation may be in the fact that these boys are also low on Factor D, which is also contrary to the reported tendency. It is, of course, believed by the layman that delinquents are quick tempered, but this is not clinically supported, nor borne out in the present findings.

D. SUMMARY

To summarize these findings, the delinquents of this sample show rebellion which finds behavioral expression in an adventurous, frivolous, carefree extraversion, with considerably less than normal anxiety. Compensation for a weakened masculine identity is accomplished by a hypermasculine facade.

Some explanation for the small but significant differences in factor means between our sample and the standardization sample was promised. It has been observed clinically for some time now that there seem to be different types of delinquents. Millar, Radman, and Sneed (5) recognized this. Two types come immediately to mind: the belligerent, hostile boy who is openly defiant of adult authority and the frequently observed withdrawn, schizoid boy who operates alone and whose actions against society are acts upon things or symbols (for example, the boy who burns down the community church). These two types would seem almost to be personality type opposites and if equal numbers of them were averaged their personality factor scores

would probably approach the population factor means. The hypothesis is offered that the existence of delinquent types has attenuated our statistics, although an analysis of Table 2 shows that 10 means are significantly different from the population upon which the test was standardized. Also, by simply looking at individual HSPQ profiles it is obvious that there is a great deal of divergence of individuals from the mean. This great divergence, coupled with significant mean differences from the normal population and statistically different standard deviations, points to the existence of delinquent types. This hypothesis is supported by clinical observations of the behavior of delinquents.

Within the BJR there is a study of delinquent typology and it is expected that this study will account for the small mean differences in this rather extensive sample.

Pending a more precise understanding of delinquent typology, it is proposed that clinicians reading individual HSPQ profiles (Forms A and B combined) for diagnostic purposes look for factors as expressed in the following formula:⁸

$$\text{DELST} = \frac{3A + C + 3(11-D) + 2(11-E) + 2(11-F) + 2H + I + 2J + 2(11-Q3) + 2(11-Q4)}{20}$$

Users of DELST should recognize that it is an approximation and is tentative, but it does give a reasonably accurate estimate on a single score that should be useful for clinical and research purposes. Each factor is weighted by rough approximation of the relative difference from the norm of that factor. Use of Forms A and B is recommended, with the combined stens being used in the formula. It must be emphasized, of course, that this formula spots the typical adolescent delinquent, but not necessarily the psychopath.

One final observation is that the two forms of the HSPQ, for some unexplained reason, measured in the same sample, bipolar opposites on Factor I. Also an analysis of Table 2 shows that Form A distinguishes significantly delinquents from nondelinquents on a number of factors, whereas Form B often fails to meet our .01 confidence limits on the same factors measured by Form A. While the sample size of Form B is considerably smaller and could account for its failure to achieve statistical significance, it is nevertheless reiterated that both forms be administered for diagnostic purposes in order

⁸ This formula has frequently been called DELST (an abbreviation for "delinquency sten") which was used as a computer memory location, and will herein be referred to by this designation.

to achieve maximum reliability. The appropriate tables for *combining* the two forms into a single sten should be used.'

REFERENCES

1. BURT, C. L. *The Young Delinquent* (3rd rev. ed.) London: Univ. London Press, 1938.
2. CATTELL, R. B., BELOFF, H., & COAN, R. W. *Handbook for the IPAT High School Personality Questionnaire (the HSPQ)*. Champaign, Ill.: Institute for Personality and Ability Testing, 1962.
3. GROSSBARD, H. Ego deficiency in delinquents. *J. Soc. Casework*, 1962, **43**, 171-178.
4. IPAT Information Bulletin # 4. Champaign, Ill.: Institute for Personality and Ability Testing, 1960.
5. MILLAR, T. P., RADMAN, R. J., & SNEED, D. Program development in a school for delinquent boys. *Soc. Casework*, 1960, **41**, 472-480.
6. PIERSON, G. R. The clinical efficacy of factorial measurement of delinquency. Unpublished manuscript, Green Hill School, Chehalis, Washington, 1962.
7. ———. Factorial changes in group psychotherapy. Paper presented at the Mental Health Research Institute, Washington, 1962.
8. TYLER, V., & KELLY, R. Cattell's HSPQ as a predictor of the behavior of institutionalized delinquents. Psychology Research Report No. 2, Fort Worden Diagnostic and Treatment Center, Port Townsend, Washington, 1962.

Psychological Services
Green Hill School
Chehalis, Washington

Psychological Services
Fort Worden School
Port Townsend, Washington

VARIABLES AFFECTING REPORT OF AWARENESS IN VERBAL CONDITIONING*¹

VA Hospital, Palo Alto, California and Stanford University

LEONARD KRASNER AND LEONARD P. ULLMANN

A. INTRODUCTION

The observation has been made that awareness should *not* be an important variable in verbal operant conditioning studies because the two phenomena don't belong together (3, 24). Operant conditioning implies research within a broad Skinnerian framework, and theoretically should *not* concern itself with such concepts as awareness which deal with inner thoughts, ideas and hypotheses. Although Skinner admits to the existence of private stimuli (23), generally such stimuli are not given systematic status among variables controlling behavior. However, most verbal conditioning investigators, perhaps because of their interest in clinical phenomena, have not been able to avoid the tempting trap of "awareness" and have sought to obtain associations between the two sets of behaviors, subjects' responsivity to verbal conditioning and subjects' postexperimental interview reports of awareness of examiners' reinforcing behavior.

Awareness is a difficult term to define (1). What is usually meant by the term in papers dealing with verbal conditioning is an *overt verbal behavior* by the subject in which he makes hypotheses or describes the experiment. Such verbal behavior may be "correct" in varying degrees: the subject states the contingency between his previous behavior and the examiner's reinforcing behavior. This statement can then be reliably rated as to its degree of accuracy. For example, in previous work (17) using a five-point scale for *degree of relative awareness*, two raters obtained 98 per cent agreement for placement of a subject's response at one of the five particular points.

For purposes of this paper awareness will be defined as an overt verbal behavior by the subject about the experiment which is rated by the examiner

* Received in the Editorial Office on April 22, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ From the Behavioral Research Laboratory, Veterans Administration Hospital, Palo Alto, California. This work was supported in part by grants M-2458 and M-6191 from the National Institutes of Health and portions of the paper were presented to the convention of the American Psychological Association, St. Louis, September, 1962.

for objective accuracy. It is important to note that at least two conceptually distinct factors are involved. The first of these is the information that the subject has available to him. This may be manipulated by the selection of subjects or by varying experimental conditions. While not using verbal conditioning, Hefferline (8) makes a very strong case for the explicit training of subjects to increase their range of bodily awareness. To quote Hefferline (8, p. 130):

If the subject reports the thumb twitch when it occurs and ceases to report it when it has not occurred, he will have become "aware" of the thumb twitch in the ordinary sense of the word—that is, he can now talk about the thumb twitch to the extent of "tacting" or "naming" it . . . there are a number of "crutches" available by which the reporting response may be facilitated. For instance, he may "watch" his response on a meter or scope, and thus be able to match the unmistakably clear visual aspects with the vaguer proprioceptive "feel," and perhaps come to say something about the behavioral context in which the response occurs.

While adequate information may be conceived as a necessary condition for subjects' accurate awareness statements, adequate information alone may not be sufficient to lead to accurate awareness statements. At a first level, the subject must be questioned (stimulated) by the examiner so that he gives answers about the experiment. Even if the information necessary is available to the subject and he has understood the examiner's question, it is still possible that he will not make the verbal responses which would lead to a categorization of "aware." The subject's personality and his relationship with the examiner may be hypothesized to be a second set of factors relevant to this particular situation. For example, Krasner and Ullmann (14) reported a case in which a paranoid patient spontaneously verbalized the correct contingency during the conditioning part of a verbal conditioning situation: he was being reinforced for "mother and mother referents." Following this verbalization, his use of the word "mother" dropped to zero. However, at the conclusion of the experiment, no "awareness" could be elicited during the routine awareness interview. Thus, although he clearly had the information, we may assume that the failure to report it was related to his personality or his relationship with the examiner.

In the first studies which investigated verbal conditioning (11) there was very little evidence of reported awareness in subjects whose verbal behavior changed significantly in the specific direction being reinforced. Awareness was approached via an interview at the end of the study in which subjects

were asked general questions as to what they thought was the purpose of the study. Subjects generally were unable to report the contingencies of the situation.

The verbal conditioning paradigm touched on the time-honored psychological problem of learning without awareness, and seemingly presented strong evidence of a laboratory demonstration of such learning (1). However, as more studies accumulated, certain methodological defects and criticisms in measuring awareness appeared which brought into doubt the interpretations of the early findings (3, 4, 6, 18, 20, 24, 25). Primarily, the interviews were considered to be *not* detailed or explicit enough to elicit the information available to the subject. They were described as unreliable and generally poor measuring instruments.

In recent studies, two major techniques are emerging in continuing these investigations. The first constitutes a refinement of the interview approach. The basic hypothesis of investigators in this group is that when one obtains what seems to be conditioning without verbalized awareness, it is due to insensitive interviewing procedures. Therefore, the experimenter must proceed to ask more probing questions until he elicits verbalization of the information he assumes must be present. For example, Levin (18) found that if only four awareness questions were used, as in most early studies, evidence for learning without awareness was obtained, whereas when more probing items were used, this evidence disappeared as responses indicating awareness of the contingencies were elicited from the subjects.

Spielberger's findings using this expanded interview technique in a series of papers (24, 25), were consistent with Eriksen's (6) conclusion that "in situations where cues and reinforcement are salient enough to produce learning, they will not escape detection by awareness." Spielberger found that only aware subjects, those who could verbalize the contingencies, conditioned in a sentence completion task.

First Tatz (27) and then Dulany (3, 4) extended this approach by demonstrating that it is possible to investigate the subject's hypothesis making aspects of "awareness." Dulany found that subjects tend to form behavioral hypotheses that are accompanied by corresponding self-instructional sets or "intentions." Dulany further demonstrated that even "wrong," but "correlated" hypotheses could, under certain circumstances, account for the mediation of the correct responses.

Investigators such as Spielberger, Dulany, and Levin are well aware of the potential weaknesses in their approach; namely, the possibility that verbalization of the correct awareness responses (information) may be suggested and

mediated by minimal cues during the interview, or that the content of the interview itself may influence the subject to look back at the task and *then* formulate the correct contingencies, thus reporting an awareness that did not exist previously. Further, some subjects less sophisticated than college students, with poorer vocabularies and vaguer conceptual categories, may not be able to verbalize adequately the difficult concepts of awareness. The implication of this view is that awareness *mediates* behavior change, and unless there is "awareness" there can be no behavior modification.

The second approach to awareness is not necessarily a denial of the Spielberger-Dulany position, but rather it avoids the semantic quicksand of "how can you *really* know whether a subject was aware at a given time" or "did he have the information." Rather than being concerned with what may be an insoluble problem, awareness is approached as a dependent variable in the verbal conditioning situation that can *in itself* either be influenced *directly* by reinforcement or *indirectly* as a function of manipulating instructional set, personality interaction, atmosphere, or the other variables which also influence conditioning.

For example, Kanfer and Martson (9) suggest ambiguity as a variable related to awareness. Lessened ambiguity leads to not only better learning, but also greater *reported* awareness. They found that the probability of verbalization of awareness can be controlled by manipulating the conditions related to increased or lessened ambiguity such as the amount of information given about the nature of the task. Their view is that "awareness" is a class of *coincidental* responses which covary with learning as a function of situational variables. Kanfer and McBrearty (10) present evidence which supports their hypotheses that the presentation of stimuli which vary sharply in similarity from alternate choices affects both the frequency of reported awareness as well as the rate of acquisition. Marston, Kanfer and McBrearty (19) conclude that awareness should be treated as "a group of verbal responses which are available in the S's repertoire and which vary in probability depending upon various stimulus conditions." Verplanck (28) recently has demonstrated that the *subject's report* of the hypotheses or "awareness" he uses in making decisions, essentially the focus of the usual "awareness interview," can be influenced by reinforcement in the same manner as other verbal behavior. In Verplanck's study, the subjects overtly verbalized the rules which governed their decisions in discrimination learning situations. Verplanck demonstrated that these verbalizations were themselves influenced by verbal reinforcement.

In a series of previously reported researches, cues about awareness have

been manipulated experimentally. Krasner, Weiss, and Ullmann (17) demonstrated that differential responsivity to verbal conditioning can be elicited from subjects by manipulating awareness cues by instructional sets. They concluded that by varying the amount of information available to the subject about the nature of the task his responsivity to verbal conditioning was affected. Consequently, the subject's behavior is a complex function of his personality, his attitudes towards the examiner, and his conception of the social influence being exerted. In an experiment by Ekman, Krasner, and Ullmann (5), half the subjects were made explicitly aware of the examiner's reinforcing behavior after operant trials, while the other half were not. Further, within each "awareness" group, the situation was structured to half the subjects as a test of their *personal problems*, while for the other subjects the experiment was structured as a measure of their *empathy*. Aware subjects for whom the experiment had been structured as a measure of personal problems showed a slight *decrease* in responsivity under reinforcement, while aware subjects for whom the experiment had been structured as a measure of *empathy* increased significantly under reinforcement. In short, using responsiveness under reinforcement as the dependent variable, instructed awareness alone did not influence the results, but a combination of instructed awareness and manipulated subject evaluation of the experiment did influence the results.

We turn from Verplanck's reinforcement of awareness statements and our own manipulations of subject's awareness to the present data on the correlates of report of awareness. The data to be presented in this report come from four verbal conditioning studies and are presented to demonstrate the function of the subject's personality and the experimental conditions in the production of reported awareness. The basic hypothesis is that reported awareness is a verbal operant behavior influenced by the same variables as any other responsivity to verbal operant conditioning.

B. PROCEDURES AND RESULTS

In the Ekman, Krasner and Ullmann experiment mentioned above (5) subjects had been assigned to an aware and an unaware condition, and within each half of these the situation had been structured as positive (*empathy*) or negative (*personal problems*). Among the subjects uninstructed as to awareness, those in the negative condition tended to report relatively greater awareness than those in the positive condition (Mann-Whitney $CR = 1.47$). In short a slightly unpleasant or threatening situation not related specifically to the examiner's behavior led to increased reported awareness.

In an experiment (29) with 64 hospitalized male psychiatric patients, only nine of the subjects reported awareness at any level greater than zero on the Krasner, Weiss, and Ullmann (17) scale. Comparing these nine cases with the 55 "unaware" subjects, it was found that the relatively aware subjects had significantly lower chronicity scores, i.e., had spent less time in mental hospitals since age 18 than the unaware subjects. It is possible to speculate that the more chronic patients had undergone extinction for verbalization of perceived interpersonal relationships. Further, it was found the aware subjects had significantly higher Welsh factor A scores than the unaware subjects; when test-taking set was partialled, the biserial correlation was .39 between Welsh A and awareness. These data were consistent with the chronicity score data; the more aware subjects were those who were more reactive and likely to seek and recognize interpersonal threat. As in the immediately preceding study, all other things being equal, a mild increase in threat or stress may increase report of awareness.

This was further illustrated with a group of advanced medical students (16). Each of three examiners conditioned 16 male medical students. Two of the examiners were male Ph.D.s who introduced themselves as "Doctor." For each of these examiners seven of 16 subjects, or 44 per cent, reported awareness in the postconditioning interview. Not one of the 16 subjects conditioned by the third examiner, a female non-Ph.D. who introduced herself as "Miss," reported awareness. Further, comparing the 14 subjects of two male examiners who reported awareness with the 18 relatively unaware subjects, there was a trend for more aware subjects to be somewhat more hostile and resistive to interpersonal influence, as measured by personality tests, than the relatively unaware subjects.

In short, in three experiments using patients, undergraduates, and medical students, there was evidence that a potentially threatening situation increased the number of subjects reporting relative awareness, and that this report was also influenced by pre-experimental personality and experience.

Finally, in order to increase the number of subjects who *would* report awareness, 80 college students were used who had been exposed to lectures, reading assignments and exercises on verbal conditioning (29). The subjects were interviewed after they had undergone one of two different experimental procedures. Those in the first condition had undergone a negative interaction with the examiner and the other half of the subjects had undergone a neutral interaction with the examiner. This was a group of subjects who both intellectually and academically should have been aware. However, 62 per cent of these subjects did not report awareness beyond level two (17)

when interviewed. The first finding with this group made use of report of awareness as the *dependent* variable to measure the effect of the experimental manipulations of induced hostility. It was found that subjects who had undergone hostility induction were less likely to report awareness than subjects who had not been made hostile. That is, 27 per cent of those in the hostile condition reported awareness that was rated three or four on the Krasner, Weiss, and Ullmann (17) scale in contrast to 47 per cent of those in the neutral condition ($t = 1.58$, df , 78, where for $p = .05$ one-tail, $t = 1.67$). This would indicate that the examiner behaving in a hostile manner elicits a behavior of withholding, of not reporting awareness, in subjects who might be presumed to possess necessary information.

A further result that bolsters the concept that failure to report awareness is a negative behavior for this group of subjects, was that in the hostile induction condition, the subjects who did not report awareness were significantly more hostile on two different test measures of hostility than those subjects who did report awareness. The hostility measures were a McReynolds Incongruity test (21) ($t = 2.19$) and the Byrne cartoons (2) ($t = 1.71$).

To obtain a systematic basis for future hypotheses, a set of four by four chi-squares were computed. The four groups—hostile *vs* neutral condition, reported *vs* not reported awareness—were double stratified against high and low hostility on the McReynolds and the Byrne, and with measures of ego strength such as the Stroop test. The picture consistently obtained was of a complex interaction in which the two conditions which were most likely to lead to a failure to report awareness were the exposure of a *hostile, poor ego-strength* individual to a hostile situation, and the exposure of a *nonhostile good ego-strength* individual to a hostile situation (15).

C. DISCUSSION

At this point, four major conclusions about the relation between awareness and conditioning seem reasonable. First, the frequency of a reinforced verbal operant can be manipulated by instructional sets which alter the amount of information about the nature of the task. Here the evidence comes from studies in which "awareness" itself was manipulated (5, 17).

Second, the report of awareness itself can be conditioned by reinforcement as other operant behaviors. Verplanck's work demonstrates this (28).

Third, the variables of which the behavior of reporting awareness is a function can be manipulated to influence the likelihood of the subject reporting awareness. As examples of the kind of complex interaction involved, the results presented in this paper indicate that the behavior of

reporting awareness is more likely to occur in a hostile individual in an ordinary social influence situation. However, if hostile individuals are exposed to a hostile or threatening situation, they become less likely to report awareness.

Fourth, both conditioning and reports of awareness may be a function of the same *sets* of personality and situational variables. The factors which would tend to alter responsivity may also tend to change frequency of reported awareness in the same manner.

Further, the problem of the relationship between verbal conditioning and awareness cannot be separated from the more general methodological problems involved in verbal conditioning studies. There are certain aspects of these studies that are directly relevant and are usually overlooked. (a) Verbal conditioning as a technique of studying the social influence process is *vastly* more complex than at first seemed apparent. The subject's behavior is *not* simply an automatic response to a reinforcing cue but rather it is a function of a large number of variables and the interaction between these variables. It is because of this complexity that reports of verbal conditioning may appear at times confused and contradictory. It is also because this complexity can be programmed into their design that verbal conditioning tasks are becoming increasingly attractive as techniques for investigating behavior influence. (b) In assessing the effects of awareness, as well as other verbal conditioning phenomena, there has usually been insufficient attention focused on certain differences in the nature of the task, the differences between a restricted, elicited operant situation such as the Taffel (26) pronoun sentence completion task and the free operant, emitted, Greenspoon word-saying (7) Salzinger (22) interview, or the Krasner, (12, 14, 17) story-telling tasks. (c) Insufficient attention has been focused on the initial contextual explanation given to the subject as to the purpose of the study and the importance or non-importance of the study to the subject. (d) Insufficient attention has been focused on the degree of involvement of the subject in the task as an important variable. (e) Individual differences among examiner effectiveness are generally ignored. If the reinforcing situation is viewed as not limited to the reinforcing cue itself but as extending to the examiner himself, and if the examiner's effectiveness can be manipulated by prestige and atmosphere variables, why should psychologists assume that all examiners are equally effective to begin with? Thus, inconsistent results can frequently be explained by the lumping together of very different tasks, subjects, and examiners under the same label of verbal conditioning.

A final note deals with the broader implications of awareness studies for those

especially interested in the kinds of personality changes that can be brought about by behavior influence techniques. The view taken by the present authors, Verplanck, Kanfer, and others, that reported awareness itself is manipulatable, has implications for investigation of social influence situations such as placebos, attitude influence, and psychotherapy. It means, among other things, that reports of awareness or "insight" in the psychotherapy situation can be controlled by the therapist. If this be so, it is further confirmation of the relationship between the therapist's value system and his theoretical orientation, and the kinds of "insights" he produces in his patients (13).

D. SUMMARY

Subjects' reported awareness in verbal conditioning experiments should be viewed in the same light as their verbal behavior during the conditioning process itself. Data from four verbal conditioning experiments were presented, and in combination with a review of the literature, led to the following conclusions:

1. The level of awareness reported by subjects is influenced by the informational cues given the subjects.
2. The verbal behavior of reporting awareness may itself be conditioned.
3. The personality of the subjects and the atmosphere of the experiment are both relevant variables in the reporting of awareness.
4. The same variables which influence conditionability influence level of reported awareness, and as such, a positive correlation between the two does not necessarily imply that awareness mediates conditionability.

REFERENCES

1. ADAMS, J. Laboratory studies of behavior without awareness. *Psychol. Bull.*, 1957, **54**, 383-405.
2. BYRNE, D. The relationship between humor and the expression of hostility. *J. Abn. & Soc. Psychol.*, 1956, **53**, 84-89.
3. DULANY, D. E., JR. Hypotheses and habits in verbal "operant conditioning." *J. Abn. & Soc. Psychol.*, 1961, **63**, 251-263.
4. ———. The place of hypotheses and intentions: An analysis of verbal control in verbal conditioning. In Eriksen, C. W. (Ed.), *Behavior and Awareness*. Durham, N.C.: Duke Univ. Press, 1962.
5. EKMAN, P., KRASNER, L., & ULLMANN, L. P. The interaction of set and awareness as determinants of response to verbal conditioning. *J. Abn. & Soc. Psychol.*, 1963, **66**, 387-389.
6. ERIKSEN, C. W. Discrimination and learning without awareness: A methodological survey and evaluation. *Psychol. Rev.*, 1960, **67**, 279-300.
7. GREENSPOON, J. The reinforcing effect of two sounds on the frequency of two responses. *Amer. J. Psychol.*, 1955, **68**, 409-416.
8. HEFFERLINE, R. F. Learning theory and clinical psychology—an eventual

- symbiosis? In Bachrach, A. J. (Ed.), *Experimental Foundations of Clinical Psychology*. New York: Basic Books, 1962.
9. KANFER, F. H., & MARSTON, A. R. Verbal conditioning, ambiguity, and psychotherapy. *Psychol. Rep.*, 1961, **2**, 461-475.
 10. KANFER, F. H., & McBREARTY, J. F. Verbal conditioning: Discrimination and awareness. *J. of Psychol.*, 1961, **52**, 115-124.
 11. KRASNER, L. Studies of the conditioning of verbal behavior. *Psychol. Bull.*, 1958, **55**, 148-170.
 12. ———. A technique of investigating the relationships between behavior cues of examiner and verbal behavior of patient. *J. Consult. Psychol.*, 1958, **22**, 364-366.
 13. ———. The therapist as a social reinforcement machine. In Strupp, H. H., & Luborsky, L. (Eds.), *Research in Psychotherapy* (Vol. 2). Washington, D.C.: American Psychological Assoc., 1962.
 14. KRASNER, L., & ULLMANN, L. P. Variables in the verbal conditioning of schizophrenic subjects. *Amer. Psychologist*, 1958, **13**, 358 (Abstract).
 15. ———. Reported awareness in verbal conditioning as a function of experimental conditions and subjects' personality. Paper presented to Western Psychol. Assoc., San Francisco, April, 1962.
 16. KRASNER, L., ULLMANN, L. P., WEISS, R. L., & COLLINS, B. J. Responsivity to verbal conditioning as a function of three different examiners. *J. Clin. Psychol.*, 1961, **17**, 411-415.
 17. KRASNER, L., WEISS, R. L., & ULLMANN, L. P. Responsivity to verbal conditioning as a function of awareness. *Psychol. Rep.*, 1961, **8**, 523-538.
 18. LEVIN, S. M. The effects of awareness on verbal conditioning. *J. Exper. Psychol.*, 1961, **61**, 67-75.
 19. MARSTON, A. R., KANFER, F. H., & McBREARTY, J. F. Stimulus discriminability in verbal conditioning. *J. of Psychol.*, 1962, **53**, 143-153.
 20. MATARAZZO, J. D., SASLOW, G., & PAREIS, E. N. Verbal conditioning of two response classes: Some methodological considerations. *J. Abn. & Soc. Psychol.*, 1960, **61**, 190-206.
 21. McREYNOLDS, P. Anxiety as related to incongruencies between values and feelings. *Psychol. Rec.*, 1958, **8**, 57-66.
 22. SALZINGER, K., & PISONI, S. Reinforcement of affect responses during the clinical interview. *J. Abn. & Soc. Psychol.*, 1958, **57**, 84-90.
 23. SKINNER, B. F. *Science and Human Behavior*. New York: Macmillan, 1953.
 24. SPIELBERGER, C. D. Awareness in verbal conditioning. In Eriksen, C. W. (Ed.), *Behavior and Awareness*. Durham, N.C.: Duke Univ. Press, 1962.
 25. SPIELBERGER, C. D., LEVIN, S. M., & SHEPARD, M. The effects of awareness and attitude toward the reinforcement on the operant conditioning of verbal behavior. *J. Personal.*, 1962, **30**, 106-121.
 26. TAFTEL, C. Anxiety and the conditioning of verbal behavior. *J. Abn. & Soc. Psychol.*, 1955, **51**, 496-501.
 27. TATZ, S. J. Symbolic activity in "learning without awareness." *Amer. J. Psychol.*, 1960, **73**, 239-247.
 28. VERPLANCK, W. S. Unaware of where's awareness. In Eriksen, C. W. (Ed.), *Behavior and Awareness*. Durham, N.C.: Duke Univ. Press, 1962.
 29. WEISS, R. L., KRASNER, L., & ULLMANN, L. P. Responsivity to verbal conditioning as a function of emotional atmosphere and pattern of reinforcement. *Psychol. Rep.*, 1960, **6**, 415-426.

VA Hospital
Palo Alto, California

THE EFFECT OF EXPERIMENTALLY-INDUCED
"AWARENESS" UPON PERFORMANCE IN FREE-
OPERANT VERBAL CONDITIONING AND
ON SUBSEQUENT TESTS OF
"AWARENESS"*¹

Ohio State University

WILL K. WEINSTEIN AND REED LAWSON

A. INTRODUCTION

A great deal of ambiguity exists in the literature concerning the question of whether verbal learning can occur without the subject being "aware," i.e., without being able to state in some fashion that he is being reinforced when he emits verbal responses of a certain generic class. In most of the relevant experiments (2, 5, 8, 10, 12, 13, 14) the method of determining awareness was some form of postexperimental interview. Numerically in the majority are studies supporting a learning-without-awareness hypothesis, but the method of postexperimental assessment of awareness varies widely in intensiveness. Levin (6), among others, has pointed out the varying results that can be obtained with different kinds of interviewing.

While it would seem to be a more direct approach to the problem, there are few studies in which degree of awareness was experimentally manipulated, viz., in which the subject was told at least something about what responses he was to learn. Only Dulany (1) has studied this extensively, with results strongly urging that the subject must be aware of the appropriate response class. A more limited kind of manipulation of awareness has appeared in three studies of the effect of making *S* aware that a reinforcer was being used to strengthen *some* unspecified verbal behavior. The results of these studies have been positive (11), negative (9), and mixed (4). Only these latter studies involved the free-operant verbal learning procedure first studied by Greenspoon (3), although it is for this situation that the strongest claims about learning without awareness have been made (e.g., 12). Further-

* Received in the Editorial Office on April 25, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ The experiment reported here was the basis for a thesis submitted by the first author in partial fulfillment of the M.A. requirements at Ohio State University. The second author is primarily responsible for the present report.

more, subjects given experimentally-manipulated degrees of awareness have not been examined by postexperimental interviews.

The present experiment, therefore, had as two of its purposes the study of the effects of both "full awareness" and "awareness of reinforcement" on (a) free-operant verbal learning, and (b) the subsequent responses of Ss during postexperimental examination for awareness. An additional feature of this research was the inclusion of two new methods of postexperimental examination: (a) having Ss pick words of the correct response class from a list, and (b) having Ss questioned in part by a person ostensibly not involved in conducting the experiment.

B. METHOD

1. *Subjects and Apparatus*

Ss were 32 female and 28 male students from sections of the introductory psychology course at Ohio State University. Ss were selected by their instructors in order to make credible the stooge interruption described later. Criterion for selection was simply the possibility of a student serving for two consecutive hours at a given time.

The experimental room was a 13 1/3 feet \times 7 feet cubicle partitioned off from a larger room. A rather large table with two chairs on opposite sides was in the middle of the room against one wall. A lamp and Standard Electric Timer were on the table, and a tape recorder and microphone were concealed in the room.

2. *Procedure During the Learning Task*

The free operant verbal learning task of Greenspoon (3) was used. E (WKW) was a graduate student with considerable training in the shaping of animal and human free-operant behavior. The correct response class chosen was plural nouns. The verbal stimulus "good" paired with a head nod was used as the reinforcer.

All Ss, upon being seated in the experimental room, were told:

This is an experiment dealing with flexibility of responding. What I want you to do is to say all the different words you can think of. Say them individually; do not count, do not use sentences, do not use phrases. Because of the nature of the experiment, it will be to your advantage not to rush. Take all the time that you need. Do you understand? All right, once again, all you are required to do is say all the different words you can think of. Say them individually; do not count, do not use sentences or phrases. Do not rush. Are you ready? Begin!

Any questions were answered by repetition of the appropriate section of the above instructions.

The 60 *Ss* were randomly assigned to one of four equal-sized groups, differing both in the nature and amount of task-relevant information received and in the use of reinforcement.

Group NIX (No Information, Reinforcement): These *Ss* received an uninterrupted 40-minute session of continuous reinforcement of the correct response.

Group PIX (Partial Information, Reinforcement): These *Ss* were given 20 minutes of continuously reinforced acquisition, after which they were stopped and given the following information:

I can now tell you that there is a correct type of response in this experiment. Every time you make a correct response, I inform you that you are correct by saying "good" and nodding my head. Do you understand? Continue!

Ss then continued to receive continuous reinforcement for the correct response for the next 20 minutes.

Group FIX (Full Information, Reinforcement): These *Ss* were also stopped after 20 minutes of the standard reinforcement procedure and told:

I can now tell you that a correct response in this experiment is any plural noun. Every time you say a plural noun, I inform you that you are correct by saying "good" and nodding my head. Do you understand? Continue!

Another 20-minute session followed.

Group CONT (Control): These *Ss* received no information and no reinforcement. Their task was to emit words, uninterrupted, for 40 minutes.

Ss' responses were tabulated on a data sheet hidden from *S's* view. Each session was tape-recorded in its entirety, without *S's* knowledge, to provide a reliability check on the confederates' reports of the disguised interview and on the experimental data records.

3. Postexperimental Checks on Awareness

a. Disguised interview. An interview by a confederate of *E*, presented to *S* merely as a fellow student, was employed in order to gain information relevant to *S's* awareness without the usual complications of *E* variables and *S's* perception of the appropriateness of revealing awareness to *E*.

At the end of the 40th minute of free-operant training, the experiment was "interrupted" by a knock on the door. The confederate entered, posing

as a student under the impression that he was supposed to be in the experiment at that particular time. *E* claimed he would have to see a secretary to straighten out the schedule and went out, leaving *S* alone with the confederate.

The stooge began to converse with *S*, engaging in small talk about the introductory psychology course, and after 1-2 minutes began to determine what *S* could (or would) tell about the experiment. Stooges never knew the experimental treatment that had been given an *S*, and their questioning was essentially the same for all cases. They explained their curiosity about the experiment by saying that they had never served in a psychology experiment before and had heard some rather wild stories. They asked: (a) "What's this thing all about, anyway?" (b) "What are you supposed to do in here?" (c) "Do you get shocked or anything?" (d) "Do you get money if you do something you're supposed to?" (e) "How do you know what to do?" (f) "What does he (*E*) do?"

The order and degree of questioning varied, depending on how much information *S* gave in answering a particular question. In no case did the confederate suggest an answer to *S* or push the questioning beyond normal discourse. Only one *S*, the last one run, gave any evidence of suspecting a connection between the stooge and *E*. Eleven *S*s refused to tell the confederate anything about the experiment; at this point *S*s had not yet been requested by *E* not to tell anyone about the experiment, but this is a standard procedure in most of the experiments in which these students serve, and may have been the basis for the refusals. In all cases, after 2-4 minutes of questioning the stooge changed back to irrelevant topics. *E* then re-entered the room and dismissed the confederate in a plausible manner.

There were two confederates, one of each sex, who were undergraduate majors (seniors) in psychology.² The male confederate, who quizzed 26 *S*s, had previously served a similar role on another experiment; the female stooge, used with 33 *S*s, had no previous experience. Stooges were assigned to *S*s solely on the basis of convenience in scheduling, and one *S* had to be run without this interview because neither confederate was available.

b. Correct response identification test. The experiment continued with the administration of a multiple-choice test that took 5-8 minutes to complete. The test consisted of a single sheet of paper with 140 words, five per line, mimeographed on it. Nearly all words were nouns, and 40 were plurals. *E*

² We wish to thank Miss Jeanne Smith and Mr. A. S. Fedoravicius for their assistance as the confederates.

simply asked *S* to follow the instructions at the top of the page, which merely directed *S* to "circle the correct words from those that appear below." It was explained that in any line there might be any number of correct responses, even none. Although many *Ss* expressed considerable dissatisfaction and confusion over the brevity of these directions, such ambiguity was considered necessary to avoid suggesting too much to *S* about the experiment before the subsequent probing was done.

c. Postexperimental interview and ad-lib test. Following completion of the identification test, *E* asked *S* three questions: (a) "First of all, did you notice anything about my behavior in this experiment?" (b) "Was there a correct type of response?" (If *S* said yes, he was asked to identify it.) (c) "Was I trying in some way to influence you to give the correct response?" (If *S* said yes, he was asked how.)

A final question was actually a separate test in itself. Each *S* was asked, "Now will you give me five correct responses other than those used by you in the course of the experiment and other than those appearing on the test?" This was called the Ad-Lib Test.

Since many *Ss* had no basis upon which to answer correctly many of these questions, the phrase "response appropriate to the experiment" often had to be substituted for "correct response."

The total experimental session lasted an average of 75 minutes.

TABLE 1
MEAN ARCSIN PERCENTAGE SCORES FOR VARIOUS LEARNING AND POSTEXPERIMENTAL MEASURES (STANDARD DEVIATIONS IN PARENTHESES)

| Measure | Group (each $n = 15$) | | PIX | FIX |
|------------------------------|------------------------|------------------|------------------|------------------|
| | CONT | NIX | | |
| Overall conditioning scores | 7.34 (6.31) | 12.97 (6.94) | 15.36 (10.65) | 39.99 (12.87) |
| Scores for 1st 5-min. block | 7.58 (5.06) | 10.37 (6.93) | 7.85 (4.63) | 12.50 (8.78) |
| Scores for last 5-min. block | 6.91 (6.18) | 17.69 (19.90) | 26.61 (28.78) | 72.33 (25.88) |
| Multiple-choice test | 16.29 (5.76) | 15.33 (6.08) | 28.87 (29.10) | 37.13 (27.60) |
| Ad-lib test | 2.30 (4.76) | 14.30 (26.20) | 35.17 (36.35) | 72.00 (37.26) |

C. RESULTS

The learning data are summarized in Figure 1 and the first three rows of Table 1. Table 1 also shows the data for the two objective post-experimental tests. A Lindquist "Type I" design (7) revealed a significant treatments

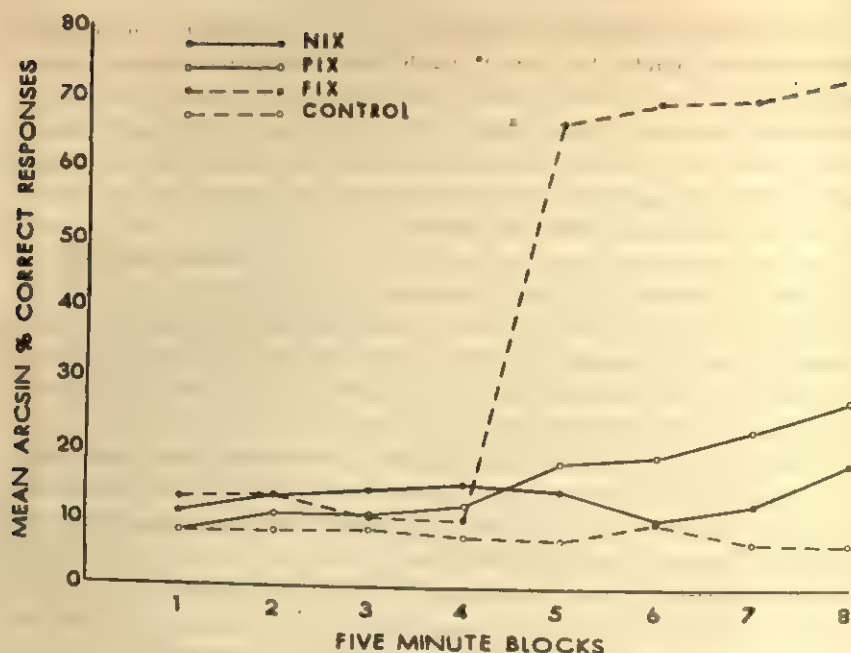


FIGURE 1

PERCENTAGE OF PLURAL NOUNS EMITTED DURING SUCCESSIVE 5-MINUTE PERIODS OF THE LEARNING SESSION

After block 4, groups FIX and PIX were briefly interrupted and given full or partial information about the nature of the experiment.

effect ($F = 340.33$, $p < .001$, 3, 56 df) and a significant blocks \times treatments interaction ($F = 23.74$, $p < .001$, 21, 392 df). A critical difference table for the overall conditioning means (first row, Table 1) showed that Group FIX had a significantly higher mean than any other group ($p < .01$ in each case), PIX was significantly higher than CONT ($p < .05$), but NIX was not reliably different from either PIX or CONT. The amount of information given seems directly related to the overall conditioning score. A similar analysis, eliminating the obviously exceptional FIX group's high means and variances, gave essentially the same results. With the revised error term, however, Group NIX just falls short of being significantly different from Group CONT. Analysis of the final 5-minute block of the conditioning session (row 3, Table 1) showed the same relationships as did the overall analysis. Analysis of the first 5-minute block (row 2, Table 1), something akin to an "operant level" test, showed no differences among the groups.

On the multiple-choice test (row 4, Table 1) the simple randomized F was significant ($3.93, p < .05, 3; 56 df$). Critical differences showed that Group FIX differed ($p < .01$) from both NIX and CONT, but no other differences were significant. If the CONT S s (who didn't know what a "correct response" was) are eliminated from the analysis, no relationships change.

On the ad-lib test, the $F = 16.38$ ($p < .001, 3, 56 df$) with all critical differences being significant ($p < .01$) except between NIX and PIX and between NIX and CONT.

The assessment of various levels of awareness via the disguised interview is summarized in Table 2. The reduction in sample sizes for various groups was due to (a) the previously mentioned refusal of 11 S s to say anything about the experiment to the stooge; (b) three cases (one each in Groups FIX, PIX and CONT) where the S was known by the confederate (and, therefore, was not interviewed); and (c) one case (CONT group) where no stooge was able to be in attendance. By this method of probing, only seven S s could be classed as fully aware, i.e., were able (or willing) to state correctly the contingency between response and reinforcement.

TABLE 2
FREQUENCIES OF VARIOUS TYPES OF AWARENESS REVEALED BY THE DISGUISED INTERVIEW

| Group | n | Aware of correct response | Aware of reinforcement | Aware of contingency |
|-------|-----|---------------------------|------------------------|----------------------|
| CONT | 9 | 0 | 0 | 0 |
| NIX | 15 | 1 | 10 | 2 |
| PIX | 10 | 2 | 8 | 2 |
| FIX | 11 | 9 | 6 | 5 |

A similar summary based on the results of E 's postexperimental interview is shown in Table 3. The data are based on the answers to the first three questions of the interview. By this method of examination, 12 S s can

TABLE 3
FREQUENCIES OF VARIOUS TYPES OF AWARENESS REVEALED BY THE DIRECT INTERVIEW

| Group (each $n = 15$) | Aware of correct response | Aware of reinforcement | Aware of contingency |
|---------------------------|---------------------------|------------------------|----------------------|
| CONT | 0 | 0 | 0 |
| NIX | 1 | 13 | 1 |
| PIX | 4 | 14 | 4 |
| FIX | 10 | 7 | 7 |

be classified as fully aware. It should be noted that one S in the FIX group who was fully aware, refused to emit plural nouns during the final 20 minutes of the learning situation. He stated, in the interview, that he was

a "contrary person" and had to be "different." Another interesting point is that only two *Ss* (both in Group PIX) were rated as fully aware by both interviewing techniques.

Analysis of variance on the overall conditioning scores of the "unaware" *Ss* in groups PIX, NIX, and CONT ($n = 11, 14$, and 15 respectively) revealed no significant difference. (Group FIX *Ss* were not included because only three could be classed as "unaware" on the basis of both interviews.)

D. DISCUSSION

Degree of experimentally-induced awareness in a free-operant verbal conditioning situation affects both the degree of conditioning obtained and the degree to which *Ss* are subsequently able to verbalize the correct contingency, to identify the reinforced class of responses, and to give examples of the reinforced class of responses. Even merely pointing out to *S* that there is a correct response being reinforced, without stating what the response class is, has a facilitative effect on all of these measures. *Ss* given continuous reinforcement without any information occupy a statistically ambiguous position midway between the partial information *Ss* and those receiving neither information nor reinforcement. When *Ss* who revealed, during one of the postexperimental interviews, awareness of the correct contingency are removed from the learning curves, no evidence of learning without awareness is obtained from this experiment.

The frailties of postexperimental assessment of awareness are nicely exemplified by this experiment. Although all of the assessment devices gave differences between groups that corresponded to the different degrees of experimentally-induced awareness, it is noteworthy that the *Ss* in the group having full information did not uniformly reveal complete awareness for any of these devices. Of particular interest is the fact that *E*, by direct questioning of a kind fairly standard in this research, could only elicit complete indication of awareness in seven of 15 *Ss* who had been told the entire purpose of the experiment less than half an hour before. Following the criteria of earlier studies, in other words, we could have shown "learning without awareness" by plotting the curve for the remaining eight *Ss* in this group (which would have been almost a replica of that for the entire FIX group shown in Figure 1).

E. SUMMARY

Learning of plural nouns in a standard Greenspoon type of experiment was studied as a function of the degree of information explicitly given to

S about the nature of the experiment. Ss told half-way through the session what was the correct response were far superior to all other Ss, but Ss told only that there was some kind of correct response being reinforced also improved their performance. Ss given reinforcement without information were not reliably different from unreinforced controls. Four different methods of assessing degree of awareness after the learning session was over gave results in accord with the amount of information explicitly given Ss, but of special interest was the fact that Ss presumably "fully aware" (because of what they had been told during the learning period) did not uniformly reveal complete awareness with any of the methods of assessment used.

REFERENCES

1. DULANY, D. E., JR. The place of hypotheses and intentions: An analysis of verbal control in verbal conditioning. In C. W. Eriksen (Ed.), *Behavior and Awareness*. Durham: Duke Univ. Press, 1962. Pp. 102-129.
2. ERIKSEN, C. W., & KEUTHE, J. L. Avoidance conditioning of verbal behavior without awareness. *J. Abn. & Soc. Psychol.*, 1956, **53**, 203-209.
3. GREENSPOON, J. The reinforcing effect of two spoken sounds on the frequency of two responses. *Amer. J. Psychol.*, 1955, **68**, 409-416.
4. KANFER, F. H. Verbal conditioning: Reinforcement schedules and experimenter influence. *Psychol. Rep.*, 1958, **4**, 443-452.
5. KRASNER, L., WEISS, R. L., & ULLMANN, L. F. Responsivity to verbal conditioning as a function of two different measures of awareness. *Amer. Psychol.*, 1959, **14**, 388.
6. LEVIN, S. M. The effects of awareness on verbal conditioning. *J. Exper. Psychol.*, 1961, **61**, 67-75.
7. LINDQUIST, E. F. *Design and Analysis of Experiments in Psychology and Education*. Boston: Houghton Mifflin, 1956.
8. NUTHMANN, A. M. Conditioning of a response class on a personality test. *J. Abn. & Soc. Psychol.*, 1957, **54**, 19-23.
9. SIDOWSKI, J. B. Influence of awareness of reinforcement on verbal conditioning. *J. Exper. Psychol.*, 1954, **48**, 355-360.
10. SPIELBERGER, C. D. The role of awareness in verbal conditioning. In C. W. Eriksen (Ed.), *Behavior and Awareness*. Durham: Duke Univ. Press, 1962. Pp. 73-101.
11. TATZ, S. J. Symbolic activity in learning without awareness. *Amer. J. Psychol.*, 1960, **73**, 239-247.
12. VERPLANCK, W. S. The operant conditioning of human motor behavior. *Psychol. Bull.*, 1956, **53**, 70-83.
13. WEISS, R. L. The influence of "set for speed" on learning without awareness. *Amer. J. Psychol.*, 1955, **68**, 425-431.
14. WILSON, W. C., & VERPLANCK, W. S. Some observations on the reinforcement of verbal operants. *Amer. J. Psychol.*, 1956, **69**, 448-451.

23 Ralsey Road South
Stamford, Connecticut

Department of Psychology
Ohio State University
Columbus 10, Ohio



THE EFFECT OF SENSORY DEPRIVATION UPON SCORES ON THE WECHSLER ADULT INTELLIGENCE SCALE*¹

Department of Psychology, Western Michigan University

MALCOLM H. ROBERTSON AND DOUGLAS J. WOLTER

A. INTRODUCTION

With few exceptions sensory deprivation research has dealt with the impairment of behavioral processes (6). The effect of deprivation on cognitive functions has been of special interest because of the practical implications that such findings might have for manned space flights. Another promising aspect has to do with the similarity between symptoms of serious emotional disorders, including varying degrees of functional intellectual deficit, and the cognitive changes in normal subjects exposed to conditions of isolation.

One of the first studies to measure the effect of deprivation on intellectual behavior was conducted by Bexton *et al.* (1954). Using tasks such as mental multiplication, anagrams, and digit span, they found that the longer the period of isolation the greater the decline in test performance. On the other hand, Vernon *et al.* (1957) failed to find any deficit in performance on rote learning after periods of deprivation as long as 72 hours. Using conditions similar to the Bexton study, Goldberger and Holt (1958) found that following eight hours of deprivation subjects showed a significant impairment on a test of logical deductions, but not on tests of arithmetic reasoning, digit span, and story recall. Heron (1961), in summarizing an experimental program conducted by the McGill group (including Bexton's study) concluded that deprivation periods of 24, 48, and 72 hours produced significant impairment on an anagram test and a word making test, a non-significant impairment on associative learning, and no deficit on a digit span test, and an analogies test.

Cohen *et al.* (1961) administered the Vocabulary subtest and part of the Digit Span, Similarities, Arithmetic, and Comprehension subtests of the WAIS to two clinical and two nonclinical subjects. The four were exposed to four hours of deprivation and were then administered the remainder of the WAIS. The subjects showed an increase on the Digit Span subtest and

* Received in the Editorial Office on May 6, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ The article is based upon a Master's thesis completed by the second author.

a decrease in scores on arithmetic reasoning, ability to abstract and generalize, and verbal reasoning.

The studies cited above differ considerably from one another on important variables such as the number and type of subjects, use of control groups, instructions given to subjects, length and type of deprivation, and response measures. Consequently, the studies neither confirm nor disconfirm one another. For example, in Cohen's study, a total of four subjects were used, two clinical and two nonclinical, and no control group was employed. In the Goldberger and Holt study, fourteen college subjects were used with no control group. In the Vernon *et al.* studies, a total of 13 experimental and 13 control subjects (college students) were used. In the McGill studies, a total of 29 experimental subjects and 27 control subjects (college students) were employed.

Regarding instructions given to subjects prior to deprivation, the Vernon *et al.* studies appeared to give the subject more information concerning the purpose and nature of the experiment than the other three studies. The instructions given to the subjects in the Goldberger and Holt study seemed to suggest that the experiment was designed to produce unusual behavioral reactions.

Though the methods of restricting sensory input were similar for the four studies, the experimental conditions of isolation were not comparable on all relevant points. Subjects in the McGill program and in Vernon's research were isolated for periods of 24 hours or longer, while in the Goldberger and Holt and the Cohen *et al.* studies eight- and four-hour deprivation periods respectively were used.

Finally, for response measures the McGill investigators and Goldberger and Holt administered a variety of intellectual tasks, while Vernon used measures designed to evaluate rote learning. Only in the experiment by Cohen was a standard intelligence scale employed.

Since the previous research has indicated that at least some types of cognitive functions may be impaired, depending on the length of deprivation, the following study was designed to assess the effects of a *short period* of isolation (three hours). The type of deprivation was similar to that used in the studies cited above. The Wechsler Adult Intelligence Scale was selected as the response measure in order to determine changes in broad intellectual functions as well as in specific abilities. An interval of three weeks between the administrations of the WAIS was considered long enough to minimize the effects of memory, and short enough to rule out any real changes in

intellectual functioning. Finally, to differentiate between the effects of deprivation and errors of measurement, a control group was employed.

B. METHOD

The sample was comprised of 20 male undergraduate students who volunteered to participate in the project. The subjects were told that the experiment was divided into two sessions, requiring a total time of approximately six hours. No information was given to the subjects as to the nature of the project. They were asked not to discuss the experiment with anyone until they were notified to do so.

The Wechsler Adult Intelligence Scale (WAIS) was administered to two subjects each day, five days a week, until all 20 subjects had been examined. The WAIS subtests were administered in the same order as in the standardization of the test. All the testing was done by one examiner.

Using a table of random numbers, one of the two subjects tested on each day was assigned to the experimental group and the other subject to the control group. The 10 control subjects were given the WAIS a second time, exactly three weeks after the first administration. Although the subjects knew that they were to participate in a second session of the experiment, they did not know what the second session would involve. The 10 experimental subjects were also given the WAIS exactly three weeks after the first administration. Before taking the WAIS a second time, each experimental subject was exposed to three hours of partial sensory deprivation and social isolation.

Each experimental subject was placed on an air mattress in a sound-deadened, lightproof room. Watches, rings, medals, glasses were taken off. Subject's clothes were loosened and shoes were removed. The subject's arms and legs were encased in a cotton-lined piece of cardboard that extended from above the elbows over the hands, and from the ankles to above the knees. Subjects also wore cotton gloves. Small holes were cut in the gloves to reduce perspiration. Subjects were told that they would remain in the room for an indefinite period of time. They were assured that they would not be hurt or frightened in any way. They were informed that a tape recorder would be turned on so that any vocalization would be recorded. They were encouraged not to terminate unless they felt that it was absolutely necessary. Only one subject requested termination before the three-hour period had elapsed.

At the end of the three hours, each subject was taken from the isolation room to an adjoining room where the WAIS was administered immediately.

After completing the WAIS, each subject was interviewed briefly to determine his reactions to the isolation conditions.

C. RESULTS

For each group, means were obtained for the subtest weighted scores and the Verbal, Performance, and Full Scale *IQ*s for both administrations of the WAIS. The principal findings on the first administration of the test are summarized in Table 1.

TABLE 1
CHANGE IN MEAN SCORES FOR EXPERIMENTAL AND CONTROL GROUPS

| | Control (<i>N</i> = 10) | Experimental (<i>N</i> = 10) | Difference |
|-----------------------|-----------------------------|----------------------------------|------------|
| Information | .9 | .2 | .7 |
| Comprehension | 2.1 | .4 | 1.7 |
| Arithmetic | .6 | .3 | .2 |
| Similarities | 1.9* | .2 | 1.7 |
| Digit span | .0 | .3 | .3 |
| Vocabulary | .5 | .1 | .6 |
| Verbal <i>IQ</i> | 6.5* | 2.7 | 3.8 |
| Digit symbol | 1.3* | .7 | 1.1 |
| Picture completion | 1.1 | 2.1* | 1.0 |
| Block design | .9 | 1.2 | .3 |
| Picture arrangement | .8 | 1.0 | .2 |
| Object assembly | 2.6 | 3.2* | .6 |
| Performance <i>IQ</i> | 9.3* | 10.9* | 1.6 |
| Full scale <i>IQ</i> | 8.4* | 6.3 | 2.1 |

Note: All changes represent increase in scores.

* Significant at the .05 level.

The control group obtained Verbal, Performance, and Full Scale *IQ*s of 109.2, 107.6, and 109 respectively. The experimental group obtained corresponding *IQ* scores of 116.9, 111.8, and 116 respectively. Although the experimental group's Verbal and Full Scale *IQ*s averaged seven points higher than the control group, the differences were not quite significant. For the experimental group, the subtest weighted scores ranged from a low of 10.4 on the Object Assembly subtest to a high of 14.7 on the Comprehension subtest. For the control group, the scores ranged from a low of 9.8 on the Object Assembly subtest to a high of 12.5 on the Comprehension subtest.

Using a *t*-test, the two groups were compared in terms of mean change on each of the 11 subtests and on the Verbal, Performance, and Full Scale *IQ*s. No significant differences were obtained.

A *t*-test was then used to determine whether the changes from the first to the second administration of the WAIS were significant. The experimental group showed a change, significant at the .05 level, on the Performance *IQ*

and on the Picture Completion and Object Assembly subtests. The control group showed a change, significant at the .05 level, on the Verbal, Performance, and Full Scale *IQ*s, and on the Digit Symbol and Similarities subtests. All changes for both groups resulted in higher scores on the second administration of the test.

D. DISCUSSION

The above findings indicated that for volunteer male college subjects, three hours of partial sensory deprivation and social isolation did not have a significant effect on their scores on the WAIS. Neither the separate *IQ* scores nor the individual subtest scores were changed significantly by the deprivation condition.

When the scores from the two administrations of the test were compared, the experimental subjects showed a significant increase on the Performance Scale *IQ*, and on the Picture Completion and Object Assembly subtests. The control subjects showed a significant increase on all three *IQ* scales as well as the similarities and Digit Symbol subtests. All changes, whether significant or not, resulted in higher scores on the second administration of the test. Such changes are not too surprising in view of the sample size and the test-retest interval of three weeks.

Since the type of subjects, sample size, tasks, instructions, and method of restricting sensory input were similar to previous investigations, the absence of any intellectual impairment can best be explained by the relatively short period of deprivation. With the exception of Cohen's study, the other investigators had used eight or more hours of isolation. Although Cohen reported a decline on some of the WAIS subtests after four hours of deprivation, in the absence of a control group the findings should be interpreted as a function of the low reliability of the subtests. The findings of the present study would certainly support this interpretation.

One final point concerns the procedure for retesting the subjects. In some of the research, an effort has been made to test subjects in the isolation room while they are still under some degree of sensory deprivation. Possibly, removing the subjects to conditions of ordinary sensory stimulation for the purpose of testing may have cancelled some of the deprivation effects in this experiment. In the McGill program, subjects were tested during deprivation as well as after they had been removed from the isolation cubicle. Some impairment occurred under both conditions. In the Vernon *et al.* studies, subjects apparently were removed from the deprivation room before testing. Although Vernon and his collaborators found no significant impairment, the

testing was limited to rote learning tasks. In the Goldberger and Holt study, subjects were examined during deprivation. Again, some tests showed impairment, others did not. In the Cohen *et al.* study, subjects were removed from the deprivation conditions before testing, and a decrease in scores on some tests was reported.

Although the issue of retesting subjects during or after deprivation deserves further consideration, the really crucial variables seem to be the length of deprivation and possibly the type of mental function being assessed.

E. SUMMARY

The purpose of the experiment was to determine the effect of sensory deprivation upon intelligence test performance. Two groups of ten subjects each were administered the WAIS. Three weeks later, the WAIS was again administered. Prior to the second administration, one group received three hours of social isolation and partial sensory deprivation. While both groups showed significant changes from the first to second administration of the test, differences between the groups were not significant.

REFERENCES

1. BEXTON, W., HERON, W., & SCOTT, T. Effects of decreased variation in the sensory environment. *Can. J. Psychol.*, 1954, **8**, 70-76.
2. COHEN, S., SILVERMAN, A., BRESSLER, B., & SHMOVANIAN, B. Problems in isolation studies. In P. Solomon, P. Kubzansky, P. Leiderman, *et al.* (Eds.), *Sensory Deprivation*. Cambridge: Harvard Univ. Press, 1961.
3. GOLDBERGER, L., & HOLT, R. Experimental interference with reality contact (perceptual isolation): Method and group results. *J. Nerv. & Ment. Dis.*, 1958, **127**, 99-112.
4. GUERTIN, W., RUBIN, A., FRANK, G., & LADD, C. Research on the Wechsler Intelligence Scales for Adults: 1955-60. *Psychol. Bull.*, 1962, **59**, 1-27.
5. HERON, W. Cognitive and physiological effects of perceptual isolation. In P. Solomon, P. Kubzansky, P. Leiderman, *et al.* (Eds.), *Sensory Deprivation*. Cambridge: Harvard Univ. Press, 1961.
6. KUBZANSKY, P. The effects of reduced environmental stimulation on human behavior: A review. In A. Biderman & H. Zimmer (Eds.), *The Manipulation of Human Behavior*. New York: Wiley, 1961.
7. VERNON, J., & MCGILL, T. The effect of sensory deprivation upon rote learning. *Amer. J. Psychol.*, 1957, **70**, 637-639.

Department of Psychology
Western Michigan University
Kalamazoo, Michigan

Oakland Child Guidance Clinic
Pontiac, Michigan

SIZE DISCRIMINATION UNDER TWO CONDITIONS OF PHOTIC INTERMITTENCY* ¹

Department of Psychology, Michigan State University

THOMAS M. NELSON, S. HOWARD BARTLEY, AND ROBERTA F. WISE

A. INTRODUCTION AND PROBLEM

Traditionally, size discrimination has been conceptualized using factors understood to be spatial rather than temporal. At the explanatory level, these factors have involved processes interrelating various portions of the retina and processes interrelating adjacent portions of other segments of the optic pathway such as cortical projection areas. Also, explanation has sometimes included "constancy" for which the underlying neural processes have not yet been delineated.

The roles temporal factors may play in spatial discriminations such as size and visual acuity are not well understood. However the results reported generally suggest that timing can effect the steepness of the brightness gradient separating adjacent areas. Helson and Fehrer (11) have shown definiteness of shape to depend upon exposure time. When exposure time is brief the confluence of object and background is ill-defined and *O* perceives an area brighter than the surround but lacking well-defined boundaries. When exposure time is sufficiently lengthened the confluence of the same areas is seen as a sharp "edge." Werner (15) has showed that the length of the null period separating two successively presented concentric targets determines whether or not the first presentation is perceived at all. Bartley (2) has demonstrated that the null periods between repeated presentations determine whether areas with sharp edges or blurred patches are produced, and that the value of the null period determines the brightness of the area itself. Fry (10) showed somewhat the same thing. Also, Bartley, Nelson and Soules (14) have shown that visual acuity will change under conditions of stimulus intermittency as a function of rate of repetition and PCF.

Processes underlying spatial discriminations involving edges have been called *contour processes* (6). Contour processes take place across retina or laterally elsewhere within the pathway. Helson and Fehrer's (11) report

* Received in the Editorial Office on May 9, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ The present paper is a part of the work being done under Research Grant NSF-G-19485 from the National Science Foundation.

indicates that contour processes lead only to areal perception when they are not continuously maintained over a long enough period to go to completion. The data of Bartley, Nelson and Soules (14) not only indicate that ongoing (longitudinal) propagations are necessary for visual acuity, but also that edges depend upon certain time characteristics of this activity, *i.e.*, ongoing (longitudinal) propagations can also interfere with the completion of transverse or lateral propagations (contour processes). They discovered that rates of intermittency which produce synchrony in the optic pathway (8) (cyclic bunching of impulse discharge alternating with lesser activity), lower visual acuity while inputs which provide for more nearly uniform type of ongoing activity *through* the retina provide more freedom for activity *across* the retina to take place. In such cases, contour processes may reach completion without being interfered with. Surges of impulse activity interspersed with little or no activity seem to preclude as fine spatial resolutions as are possible with a more uniform flow.

One may infer that size discrimination will be influenced when intermittency characteristics of photic input are properly varied since the size of two-dimensional objects is defined by the location of edges.

The present paper reports effects of two different conditions of intermittent stimulations on perceived size. One condition is designed to produce nearly uniform longitudinal activity in the optic pathway, the second to produce highly phasic or synchronous activity.

It is expected that some forms of photic intermittency will produce neural effects leading *O* to attribute the edge to a position nearer the center of the object seen than will others. If no change in perceived size occurs as a result of manipulating synchrony of activity in the optic pathway, one may conclude that conditions used to vary brightness and visual acuity have no effect upon edge localization. If perceived size either increases or decreases under such conditions then one may conclude that *temporal* distribution of *longitudinal* propagation in the pathway affects the location of edges.

B. METHOD

1. Apparatus

The investigation was conducted in a darkened room 8×30 feet. At one end of the room *O* was seated at a chinrest fixed to a table supporting an upright black screen 30×40 inches. *O* observed monocularly (right eye) through a peephole in the screen. Toward the other end a lamphouse was suspended containing four incandescent lamps. The side of the lamphouse

facing *O* was made of opal glass and circle cutouts defining the target were placed against this surface. The radiant intensity of this surface was 1102 c/ft².

Seventeen circular targets were used. These varied in diameter from 3.125" to 5.250", increasing in uniform steps of .125". At the single observation distance used, the visual angle varied between 38' and 1°2'. All targets were therefore foveal. The complete range of values used is listed in Table 1.

An episotister provided two conditions of stimulus intermittency, namely, 10 cps and 48 cps with a PCF (pulse-to-cycle fraction²) of .25. The episotister was mounted just beyond the peephole. This arrangement kept onset and termination of pulses abrupt since the diameter of the rotating disk was large.

2. Observers

Fourteen members of an advanced undergraduate laboratory class served as *Os*. Visual acuities of all were normal or corrected to normal.

3. Procedure

O was seated at the table and given pencil and paper. He was told that he was to draw the size of the circular targets to be seen through the peephole. Practice in making drawn reproductions was given. It consisted in drawing the size of a white circular disk from a number of viewing distances providing visual angles within the range subtended by the experimental targets. Since the visual angles involved were small, all *Os* introduced a factor of magnification into their drawings. This varied considerably from *O* to *O*. No attempt was made to train *O* to use a given range of size reproductions. Instead practice in responding was continued until individually consistent sets of drawings occurred.

Intermittency conditions were run separately. Targets were presented in random order under the first occurring condition until each target had been reproduced five times, and then the intermittency condition was changed. The 10 cps condition preceded the 48 cps condition for seven *Os* and the 48 cps condition preceded the 10 cps condition for the other seven.

Target exposure was limited to 10 seconds by a Hunter Timer wired in circuit with the lamps. Response was made only after exposure to the target for the full period of presentation. Response was evaluated by measuring the diameter to the nearest millimeter. Since most drawings were somewhat irregular in outline, the broadest internal dimension was defined as the diameter of the reproduction.

² In older terminology this is known as the "light-dark-ratio" (LDR).

C. RESULTS

Table 1 presents the mean responses of the 14 *O*s. As can be seen, all 17 means are smaller for the intermittency producing brightness enhancement (10 cps, PCF .25) than for that approaching Talbot level (48 cps, PCF .25).

TABLE 1
VISUAL ANGLES AND MEAN REPRODUCTIONS OF THE INTERMITTENT TARGETS
(Each mean is based upon an *N* of 70 reproductions)

| Target | Visual angle | Mean reproduction in cm | |
|--------|--------------|-------------------------|--------|
| | | 10 cps | 48 cps |
| 1 | 38' | 12.29 | 15.16 |
| 2 | 39' | 13.69 | 14.46 |
| 3 | 41' | 13.56 | 15.75 |
| 4 | 42' | 14.39 | 16.26 |
| 5 | 44' | 15.10 | 16.79 |
| 6 | 45' | 14.96 | 16.55 |
| 7 | 47' | 15.64 | 17.38 |
| 8 | 48' | 15.09 | 17.79 |
| 9 | 50' | 16.94 | 17.46 |
| 10 | 51' | 14.85 | 17.82 |
| 11 | 53' | 17.88 | 19.56 |
| 12 | 54' | 18.37 | 19.36 |
| 13 | 56' | 18.06 | 19.48 |
| 14 | 57' | 19.27 | 20.26 |
| 15 | 59' | 19.01 | 21.21 |
| 16 | 1°0' | 19.32 | 21.26 |
| 17 | 1°2' | 19.47 | 21.06 |

The size difference from one condition to another was not large for most subjects but nevertheless consistently in the direction of differences in the average data given in Table 1. Since all *O*s served under both conditions Wilcoxon's Test for Paired Replicates was used to estimate the significance of differences. Comparison of average response from single individuals resulted in $T = 19$ which is significant at better than the five per cent level of significance using a two tailed test. Thus response differences between conditions seem likely not to have resulted from mere accident.

D. DISCUSSION

Although the brightness index $\frac{\text{intensity of steady target}}{\text{intensity of intermittent target}}$ involved in the brightness match increased approximately 2-1/4-fold from the 48 cps to the 10 cps condition, the reproduced size decreased approximately 10 per cent. The percentage decrease falls near the median value in the range of percentage visual acuity decreases reported in the previous study (14) and

suggests that the size changes are associated with the circumstances reducing visual acuity. Both size and visual acuity changes are much smaller than those changes associated with brightness, however. The difference between the magnitude of brightness changes and edge changes is so large as to make it seem that altering the temporal distribution of input is less effective in shifting edge than in altering brightness in conditions such as those investigated. This is, of course, to be expected.

If brightness increases had been brought about by intensity increases and not by the forms of stimulus intermittency used, factors of "irradiation" (16) and stray illumination (1, 9) might lead one to expect the brighter of two targets on a dark background to appear larger. The present study, however, shows size to vary *inversely* with brightness increases induced by a reduction in rate of stimulus intermittency. The effect of increasing brightness by the temporal means used here is opposite in direction to that to be expected from intensity increases. Significantly, opposing effects from the two means of increasing brightness were also found in the case of visual acuity. Visual acuity decreased with brightness increases induced by stimulus intermittency and increased with brightness increases induced by intensity increases.

As said before the intermittency inputs in the experiment elicited different temporal patterns of activity in the optic pathway (4, 5, 7). The rate of 48 cps may be presumed to provide a summed (mass) record indicating a nearly uniform level of neural activity and the rate of 10 cps with a PCF of .25 produce neural activity having sharply phasic components (2). The smaller sizes are associated with bunched distributions of discharge presumably because successive inputs are sufficiently massed not to allow contour processes set up from their predecessors to go to completion. Inhibition occurs as the process is unfolding and since edges apparently do not form everywhere at once but move outward from the center as contour processes move to completion, when contour processes are interrupted before completion the edge is attributed to a locale nearer the center.

This hypothesis seems to tally well with what is already known. In addition to the work already cited one might cite the qualitative changes observed in the flicker-fusion studies (3). Rates of intermittency just below fusion produce what is known as residual flicker. This appears near the final edge of the field as a fine *very slow* flicker. It suggests that disruption of contour processes is minimal under higher rates of intermittency. In contrast, as rate is lowered, flicker becomes more vigorous, paradoxically rapid, and involves the more central portions of the target. This is particularly noticeable with PCF in the neighborhood of .25. Finally, all portions of the field including

the center dissipate and re-form when rate of alternation is very slow. At this point we are probably dealing with pulses separated by null periods so long that no effect from previous stimulation occurs, *i.e.*, with "isolated pulse" phenomena (12, 13).

The responses elicited from *O* depend, of course, upon *E*'s instructions. In the present case diameters of the discs were dealt with and not impressions of area. Had size been defined volumetrically the relation between "size" and intermittency might have been otherwise. The reason for this is the following. The 48 cps or fusion condition produces an appearance of a single flat surface having well-defined edges while the 10 cps condition produces an appearance having a three-dimensional characteristic. Under the slower rate *O* sees the side extremities of the field located back of a bright center. It might be presumed that the 10 cps condition appearance exceeds the more rapid in apparent surface area because of its three-dimensional character. A quantitative description of both these changes is desirable.

E. SUMMARY

A series of 17 circular targets were viewed under 2 conditions of photic intermittency. Reproductions of the diameter of the disk were reliably smaller under the 10 cps-PCF .25 condition than under the 48 cps-PCF .25 condition. Two quite different patterns of ongoing neural activity within the channels of the optic pathway are associated with the two conditions of input and these seem related to the perceptual outcome. It is suggested that interruptions of the contour processes brought about the results and that the smaller size resulting under the inhibiting condition occurs because edges form first at the center of the discriminated object and move outward.

REFERENCES

1. BARTLEY, S. H. The comparative distribution of light in the stimulus and on the retina. *J. Comp. Psychol.*, 1935, **19**, 149-154.
2. ———. Some effects of intermittent photic stimulation. *J. Exper. Psychol.*, 1939, **25**, 462-480.
3. ———. *Vision: A Study of Its Basis*. New York: Van Nostrand, 1941.
4. ———. Visual sensation and its dependence on the neurophysiology of the optic pathway. *Biol. Sympos.*, 1942, **7**, 87-106.
5. ———. Some facts and concepts regarding the neurophysiology of the optic pathway. *A.M.A. Arch. Ophthalmol.*, Part II, 1958, **60**, 775-791.
6. ———. *Principles of Perception*. New York: Harper, 1958.
7. BARTLEY, S. H., & NELSON, T. M. Some relations between sensory end result and neural activity in the optic pathway. *J. of Psychol.*, 1963, **55**, 121-143.
8. BISHOP, G. H. Cyclic changes in excitability of the optic pathway of rabbit. *Amer. J. Physiol.*, 1933, **103**, 213-224.

9. BOYNTON, R. M. Stray light and the human electroretinogram. *J. Opt. Soc. Amer.*, 1953, **43**, 442-449.
10. FRY, G. A. Depression of the activity aroused by a flash of light by applying a second flash immediately afterwards to the adjacent areas of the retina. *Amer. J. Physiol.*, 1934, **108**, 701-707.
11. HELSON, H., & FEHRER, E. V. The role of form in perception. *Amer. J. Psychol.*, 1932, **44**, 79-102.
12. NELSON, T. M., BARTLEY, S. H., & JEWELL, R. M. Effects upon brightness produced by varying the length of the null interval separating successive "single" pulses: Sensory implications of the Alternation of Response Theory, I. *J. of Psychol.*, 1963, **56**, 99-106.
13. ———. Brightness changes associated with variation in the number of pulses in a repetitive photic train: Sensory implications of the Alternation of Response Theory, II. *J. of Psychol.*, 1963, **56**, 107-113.
14. NELSON, T. M., BARTLEY, S. H., & SOULES, E. M. Visual acuity under conditions of intermittent illumination productive of paradoxical brightness. *J. of Psychol.*, 1963, **55**, 153-163.
15. WERNER, H. Studies in contour: I. Qualitative analyses. *Amer. J. Psychol.*, 1935, **47**, 40-64.
16. WILCOX, W. W. The basis of dependence of visual acuity on illumination. *Proc. Natl. Acad. Sci.*, 1932, **18**, 47-56.

Department of Psychology
Michigan State University
East Lansing, Michigan

NOTES CONCERNING "NONMOTION" RESPONSES WHILE VIEWING ROTATING STIMULUS OBJECTS*

Perception Laboratory, Veterans Administration Hospital, Bedford, Massachusetts

THOMAS MULHOLLAND

When normal subjects viewed a constantly rotating stimulus object, a variety of illusory motions were perceived as well as the actual motion. In addition to the perceptual changes, other behaviors occurred which illustrated the principle of "organismic" response whereby many integrated response systems are simultaneously activated. The reactions of subjects to moving stimuli which refer to perceptual motion have been reported elsewhere and are presented only briefly here (1, 2, 3, 4). However, much less attention has been paid to the other classes of responses to moving stimuli. A qualitative description of a representative sample of such responses is presented here. These observations were selected from the records of a long series of studies. Since the "nonmotion" responses were not the main focus of these studies (3, 4, 5) they were recorded only when they were manifest enough to be noted as incidental phenomena.

A. MOTION RESPONSES

While viewing a small dimly illuminated stimulus object, which was turning at a constant speed clockwise, subjects reported a variety of illusions as well as the actual motion. These perceived motions are of two kinds: three-dimensional and two-dimensional motion. Three-dimensional motions include the perception of the actual rotation clockwise motion and perception of illusions of counterclockwise rotation and oscillation. When the oscillation illusion occurs the object seems not to rotate but rather to swing back and forth. Two-dimensional motions (illusions) are (a) expansion contraction—the object seems not to turn but rather alternately to expand and contract in two dimensions (some subjects call this "bouncing"); (b) reversal of figure and background—this is similar to (a) but the object is, for instance, seen as a hole in a black curtain which alternately opens and closes. Sometimes subjects reported no motion; rather they perceived only a light blinking on and off. This was a rare illusion.

* Received in the Editorial Office on May 9, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

B. NONMOTION RESPONSES

1. *Language*

The language used by the subjects to describe the rotation motions was usually concise and abstract. This was in contrast to the more expanded and concrete descriptions of two-dimensional motion. Usually when reporting the rotation motions almost all subjects described them as "rotating clockwise," "turning clockwise," "going around." Often descriptions of two-dimensional motion were, e.g., "expanding and contracting like a balloon," "a curtain opening and closing over a white screen," "like a dancer throwing out her arms," "like a thermometer with two black dots running back and forth." A more elliptical definition also used was, "expanding-contracting" (often accompanied by hand gestures). No subject gave a description which was an abstract geometrical description such as "deformation having a variable linear velocity in the frontal plane," though many of the educated Ss tested possessed such language skill.

2. *Expressions of Affect*

a. Relative to the perceived motions. No expressions of affect relative to the rotary motions were noted. However, such expressions occurred relative to the two-dimensional motions, e.g., "expanding and contracting like a balloon, *very cute*"; "bouncing (two-dimensional) *was more pleasing*." Exclamations or expressions of surprise often occurred when the oscillation or two-dimensional illusions were first perceived.

b. Relative to the situation. "Visual field feels very primitive; disturbing if it gets that way." "Strange experience with the background—the background tries to draw your attention from the figure." "Most unstable (visual field). Sort of disconcerting to see your world melting." Another kind of feeling was that of being hypnotized, e.g., "Gets very fuzzy toward the end. Felt not sleepy but sort of hypnotized. Hypnagogic space—everything seems to be mushy, fuzzy. I forgot mundane apparatus."

3. *Awareness of Voluntary Control*

During viewing some Ss would spontaneously discover they could control the occurrence of certain perceived motions "at will." [Studies have shown that most Ss can control the occurrence of the perceived motions by means of a voluntary set induced by instructions (5).] This discovery was never reported with negative feeling, rather it was a pleasant experience.

4. *Beliefs Concerning the Causes of the Perceived Motions*

While most normals and many psychotics spoke of an internal cause related to themselves, e.g., "your eyes play tricks on you," many externalized causes and reacted to the perceived motions as if they were real. This externalization of cause was especially marked in the case of some of the paranoid schizophrenics. In two experiments, 20 schizophrenics and 20 normals were interviewed¹ concerning the causes of the perceived motions. Three questions were asked during the interview: What causes the various motions? What did you do when you tried to control the motions? Were you able to control the motions?

Summaries of some representative protocols are presented below.

P.M., Normal:

It seems if you really stared at the card hard—that's when it seemed to have those different appearances. It looked pretty good to me that it did turn counterclockwise—I don't think you could do that just by looking at it. I thought he was doing it somehow—I don't know—I think he can turn the thing around as far as clockwise and counterclockwise—I think he controls that—but as for contracting and expanding and waving I think those are illusions.

F.K., Normal:

I thought the other fellow was controlling it with push buttons. Should be a little motor—a switch, sort of cuts in, makes it go one way, then another.

C.N., Schizophrenic (in good contact):

Well, I can't say, but something about the waves that can make it move, how it can move—electricity and waves. The waves controlled it—the waves make it move around, like the wind—wind, waves, well, the electricity could do it maybe but the wind comes first. I can't say if I'm right or wrong.

J.B., Schizophrenic:

Sexy—the little black spot reminded me of a girl's hymen, but I've never seen one. The opening and closing reminded me of the muscles that contract a woman's vagina—something sexy about the doctor's voice and in the words opening and closing. I saw a white cross—in the same test I saw a crucifix cross—I also saw a window shade and an iron post as high as this table they have in front of the ABC Station at Leverett Station where the boats go through.

I thought the doctor was doing it through his vital organs. I thought

¹ By Arnold Gofstein, Ph.D., whose assistance is appreciatively acknowledged.

he had a woman's instead of a vagina. A truthful answer—I don't know if I'm smart and truthful, but a truthful answer saved me in World War II. I had a suspicious thought that it was all being done by some power the doctor had. I don't know what time elapsed but I had the thought that maybe the cord and the plug did it. I saw the doctor go to the black box, to the door—he was trying something. If I had to testify I saw a wire, I'm not sure I saw a wire leading to the plug—I thought I saw a wire and when you think you don't know. The second and third, I had a suppliant thought, I couldn't get hypnotized by the thing going around and the monotone or soothing voice of the doctor. I had a medical book and it said that hypnosis is brought about by the patient coming willingly and letting the doctor suggest. I tried to jam the opening and closing with my fists but I couldn't. They trip football players up that way watching a football game. I had a vague fear that I was going to fall into a hypnotic state—I made defensive gestures with my feet and elbows to keep from being hypnotized. I can see every revolution but my eyes became strained. Hypnosis fear lasted only two seconds then I caught wise and became sarcastic, only for a moment or two. Then I will cooperate with the doctor only in watching the machine and not in creaming off.

J.B., Schizophrenic (in poor contact):

I think there are two magnets on one side of the camp shaft—as soon as you put on the current the connection to the camp shaft kept the electrical energy revolving and the impulses made it go around and revolve. I think that power did it on the bottom—the power tacit inside me—he gets that power as soon as that rhythm of deceiving goes in and out, two people together in a romantic way—you have a combination of two powers having a time of peace (piece?) for themselves—I can see my face in it—electrical impulses went from my chewing gum right into the screen—electrical impulses like three people dancing around—it comes right through the chewing gum, around and around, left and right—I leaned way back after—some prankster—some flight happened to somebody and they didn't tell it wasn't good enough.

It is interesting that these extensive externalizations and sexual fantasies are much like those obtained with “projective” tests.

5. Reality Testing Behaviors

Very often, at the end of the experimental session Ss would try to find out how the object “really moved.” This reality testing usually began with a question to *E*; if his answer were not enough, *S* would often walk over to the apparatus and either try to peer in or lift up the box. Many Ss asked if others “saw it the same way.” Some Ss refused to accept the idea that the motions were illusory despite their own “successful” reality-testing

behavior (looking in the box) which led to knowledge that the object was constantly rotating. It was the author's impression that those who maintained that the motions were caused by external agents (not illusions) were those who did not usually exhibit reality testing behavior at the end of the session.

6. *Group Reactions*

Pressures toward consensual agreement were clearly developed during classroom demonstrations where the group viewed the object and reported their percepts aloud. At first there was 100 per cent agreement; everybody saw rotation clockwise. However, as the illusions began to occur at different times for different *Ss*, expressions of surprise, frowns, and laughter occurred more and more often. At the conclusion of the demonstrations a group discussion was initiated with a goal to decide the "real motion of the object." Frequently, to the discomfiture of the group, agreement could not be reached. In these classroom demonstrations some *Ss* reported that finding out that their vision was subject to deviation from reality or group belief was a distinctly unpleasant experience.

C. SUMMARY AND DISCUSSION

While viewing a constantly rotating stimulus object the subjects reported that various kinds of motions were perceived. At the same time they often exhibited different forms of descriptive language, affect expression, externalization of causality, and reality-testing behavior. These phenomena point up the way in which the whole personality is involved in responding to a "neutral" stimulus.

The different perceived motions may reflect different levels of perceptual integration and the other response systems related to these perceptions may vary with the level of perceptual integration. Illusory two-dimensional motion is much more frequently described in concrete, metaphorical language with more expressions of feeling and affect than is the case with the rotation motions. In Werner's terms, there is more concrete language, less differentiation of feeling and perceiving, and less perceptual integration when *S* perceives two-dimensional motion compared to perceiving rotary motions (6).

Externalizations of causes of perceived motions were often seen and in psychotics were definite and extensive. The externalizations and sexual fantasies evoked were similar to those typically seen with projective tests. Some typical reality testing behaviors and group reactions to the moving stimuli were described.

REFERENCES

1. AMES, A. Visual perception and the rotating trapezoidal window. *Psychol. Monog.*, 1951, **65** No. 324, 3-27.
2. MILES, W. Movement interpretations of the silhouette of a revolving fan. *Amer. J. Psychol.*, 1931, **43**, 392-405.
3. MULHOLLAND, T. Motion perceived while viewing rotating stimulus objects. *Amer. J. Psychol.*, 1956, **69**, 96-99.
4. ———. The swinging disk illusion. *Amer. J. Psychol.*, 1958, **71**, 375-382.
5. ———. Instructional sets and motion perceived while viewing rotating stimulus objects. *J. of Psychol.*, 1963, **56**, 233-237.
6. WERNER, H. *Comparative Psychology of Mental Development*. Chicago: Follet, 1948.

Perception Laboratory
Veterans Administration Hospital
200 Springs Road
Bedford, Massachusetts

INSTRUCTIONAL SETS AND MOTION PERCEIVED WHILE VIEWING ROTATING STIMULUS OBJECTS*

Veterans Administration Hospital, Bedford, Massachusetts

THOMAS MULHOLLAND

A. PURPOSE

When *S* views a small, constantly rotating stimulus object he experiences a variety of perceived motions (3, 5) similar to those reported while viewing a shadow projection of a three-dimensional object (2, 6). The occurrence of these veridical and less veridical motions varies with changes of stimulus size (1); shape (1, 3, 5); speed of rotation (3); stimulus background brightness contrast (4). Other variables mentioned include "past experience" (1) and "suggestion" (2).

This paper is concerned with the effect of instructional sets on the occurrence of the various perceived motions. Inferentially this refers to the efficacy of the perceiver's control over his own perceptual content in accordance with instructions. Two kinds of sets were studied: (a) continuing sets (300-600 seconds) and (b) shifting sets (75-100 seconds).

B. APPARATUS AND PROCEDURE

A white card two inches square was constantly rotated (33.3 r.p.m.) in front of a flat black background. The background and stimulus card were illuminated by a small lamp operating at reduced voltage. The lamp was so arranged above and in front of the rotating object that no shadows were visible on the background. Illumination on the background ranged from 0.05 to 0.125 foot candles depending on conditions. The *S*, from a distance of nine feet, viewed the rotating object through a 3×4 inch aperture in the front of a large black box which enclosed the display.

During the experiments *S* viewed the object and reported his perceptions at approximately 5-second intervals in response to an audible click signal. Sets were induced by instruction: (a) Open-ended set (OE)—*S* was told to view the object and report what he saw; (b) Rotation clockwise (RCW)—similar to (a) except that *S* was told to "try to see only the rotation clock-

* Received in the Editorial Office on May 10, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

wise motion"; (c) Two-dimensional (TD)—same as (b) except *S* was told to "try to see only two-dimensional motion" (an illusion).

The data were frequency counts of the various kinds of perceived motions reported: *rcw.*—veridical rotation clockwise; *rccw.*—rotation counterclockwise; *osc.*—oscillatory motion (1, 2, 3); *td.*—two-dimensional motion.

The difference between set conditions was tested using Friedman Chi Square of Ranks or the Mann-Whitney *U* test. All statements of significance refer to *P* less than 0.05. (Since the effects of set were seen in each of six experiments with no reversals of direction the statistical treatment is somewhat redundant.) Subjects were hospital aids, students, staff members and tranquilized schizophrenics.

C. RESULTS

1. Continuing Sets

a. Experiment 1. Five men, who had previously experienced the perceived motion phenomena were tested under each of nine conditions (10-minute duration) presented at random. These were combinations of three instructional sets (OE, RCW and TD) with three levels of display illumination (dim— < 0.025 foot candles, brighter— 0.025 foot candles, and bright— 0.125 foot candles). One hundred twenty verbal reports were obtained for each *S* under each condition making a total of 5400 reports.

As illumination increased and with RCW set, *rcw.* was reported significantly more often, *td.* significantly less often compared to decrease of illumination and TD set. See Table 1.

b. Experiment 2. Ten men received the fixed sequence of OE, RCW, RCW, OE sets. Each set was of 5-min. duration. Sixty reports were obtained for each *S* during each condition making a total of 2400 reports. As in Ex-

TABLE 1
MEAN FREQUENCY OF REPORTS OF PERCEIVED MOTION DURING CONTINUING SET-CONDITIONS
(Experiment 1)

| Perceived motions | Instruction-sets | | | | | | | | |
|-------------------|---------------------|-----|----|---------------------|----|----|---------------------|----|----|
| | RCW | | | OE | | | TD | | |
| | Brighter Medium Dim | | | Brighter Medium Dim | | | Brighter Medium Dim | | |
| <i>rcw.</i> | 115 | 101 | 85 | 89 | 88 | 72 | 72 | 60 | 28 |
| <i>rccw.</i> | 2 | 8 | 8 | 7 | 7 | 12 | 4 | 3 | 5 |
| <i>osc.</i> | 3 | 7 | 13 | 15 | 20 | 17 | 12 | 14 | 8 |
| <i>td.</i> | 0 | 4 | 12 | 8 | 5 | 15 | 33 | 41 | 75 |
| Other | 0 | 0 | 1 | 0 | 0 | 5 | 0 | 2 | 3 |

periment 1, more reports *rcw.* were given during RCW set compared to OE set. The difference between *rcw.* and *td.* reports during RCW set was significantly greater than that obtained during OE conditions. See Table 2.

TABLE 2
MEAN FREQUENCY OF REPORTS OF PERCEIVED MOTION DURING CONTINUING SET-CONDITIONS
(Experiment 2)

| Perceived motions | Instruction-sets | |
|-------------------|------------------|-----|
| | OE | RCW |
| <i>rcw.</i> | 35 | 41 |
| <i>rcw.</i> | 5 | 5 |
| <i>osc.</i> | 8 | 5 |
| <i>td.</i> | 12 | 8 |

c. *Experiment 3.* Ten normal men and 10 tranquilized schizophrenics were tested. For the first five minutes (pre-set) all Ss received OE set. For the next five minutes, half of each diagnostic group received RCW set, half TD set. For the last five minutes (post-set) all Ss again received OE set. Each S gave 60 reports during each condition making a total of 3600 reports.

No evident difference between normals and tranquilized schizophrenics was found, consequently, their data were pooled. The RCW set group gave significantly more reports of *rcw.*, significantly less reports of *td.*, during the set activation compared to the TD set group. This effect was confined to the first series of 20 reports during the set activation. See Table 3.

TABLE 3
MEAN FREQUENCY OF REPORTS OF PERCEIVED MOTION DURING CONTINUING SET-CONDITIONS
(Experiment 3)

| Perceived motions | Instruction groups | Pre-set | | | Set-activation Interval of 1/3 viewing time | | | Post-set | | |
|-------------------|--------------------|---------|----|-----|--|----|-----|----------|----|-----|
| | | I | II | III | I | II | III | I | II | III |
| | | | | | | | | | | |
| <i>rcw.</i> | Set RCW | 12 | 6 | 6 | 15 | 11 | 8 | 13 | 10 | 11 |
| | Set TD | 13 | 10 | 8 | *10 | 10 | 10 | 12 | 11 | 11 |
| <i>rcw.</i> | Set RCW | 0 | 2 | 2 | 0 | 0 | 1 | 0 | 1 | 0 |
| | Set TD | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| <i>osc.</i> | Set RCW | 1 | 3 | 3 | 1 | 2 | 4 | 2 | 2 | 3 |
| | Set TD | 2 | 2 | 2 | 0 | 1 | 1 | 1 | 1 | 2 |
| <i>td.</i> | Set RCW | 6 | 9 | 8 | 3 | 6 | 6 | 5 | 8 | 5 |
| | Set TD | 4 | 6 | 9 | *9 | 8 | 9 | 6 | 7 | 6 |

* Set RCW group vs set TD group: Mann-Whitney 'U' significant at or beyond 0.05 level.

2. Alternating Sets

In these experiments *S* shifted his set on a prearranged signal from one kind to another during the viewing period. The signal to shift sets was so given between verbal reports that no interruption in the reporting occurred.

a. *Experiment 4.* Five men, experienced with the perceived motion phenomena, were tested for 750 seconds. During this period *S* shifted his set from OE to RCW, back to OE, etc., on signal, every 75 seconds. During each set, *S* gave 15 reports for a group total of 750 reports. After the first 150 seconds significantly more reports of *rcw.*, less of *td.*, occurred during pooled RCW sets compared to pooled OE set. See Table 4.

TABLE 4
MEAN FREQUENCY OF REPORTS OF PERCEIVED MOTIONS DURING
ALTERNATING SET CONDITIONS
(Experiment 4)

| Perceived motions | Instructional sets | | | | | | | | | |
|-------------------|--------------------|-----|----|-----|----|-----|----|-----|----|-----|
| | OE | RCW | OE | RCW | OE | RCW | OE | RCW | OE | RCW |
| <i>rcw.</i> | 13 | 11 | 7 | 11 | 6 | 11 | 6 | 12 | 6 | 11 |
| <i>rcw.</i> | 2 | 3 | 4 | 2 | 5 | 3 | 4 | 3 | 3 | 3 |
| <i>osc.</i> | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 3 | 0 |
| <i>td.</i> | 0 | 1 | 2 | 1 | 2 | 0 | 3 | 0 | 2 | 1 |

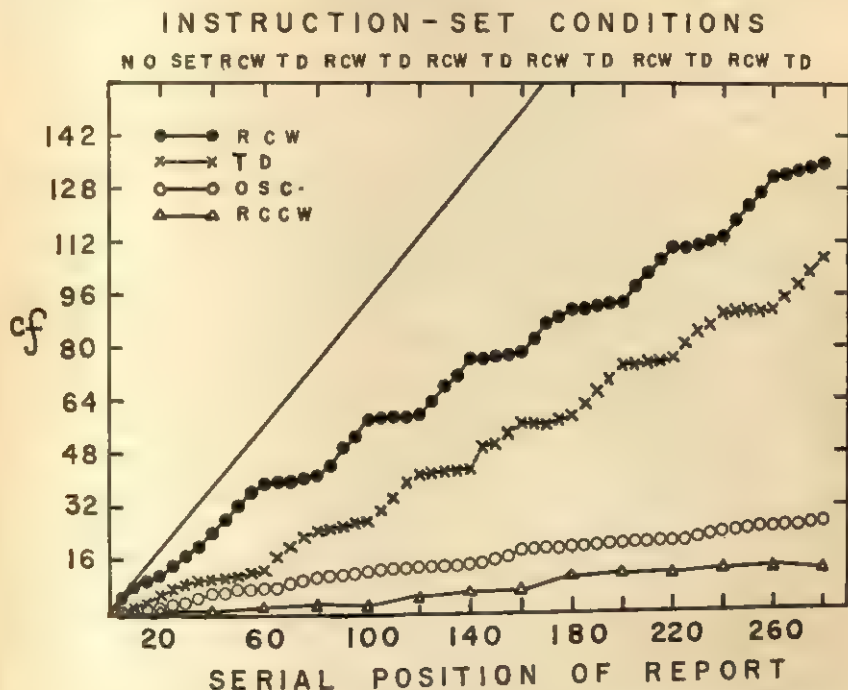
b. *Experiment 5.* This was the same as Experiment 4 except that *S* shifted between OE and TD sets. More reports of *td.* and less of *rcw.* were obtained during TD sets compared to OE sets but this result was not significant. See Table 5.

TABLE 5
MEAN FREQUENCY OF REPORTS OF PERCEIVED MOTIONS DURING
ALTERNATING SET CONDITIONS
(Experiment 5)

| Perceived motions | Instructional sets | | | | | | | | | |
|-------------------|--------------------|----|----|----|----|----|----|----|----|----|
| | OE | TD | OE | TD | OE | TD | OE | TD | OE | TD |
| <i>rcw.</i> | 13 | 7 | 8 | 4 | 8 | 4 | 9 | 4 | 9 | 3 |
| <i>rcw.</i> | 1 | 1 | 2 | 1 | 4 | 1 | 1 | 2 | 2 | 2 |
| <i>osc.</i> | 0 | 1 | 3 | 2 | 1 | 0 | 0 | 0 | 1 | 0 |
| <i>td.</i> | 1 | 6 | 2 | 8 | 1 | 10 | 3 | 9 | 2 | 7 |

c. *Experiment 6.* Five men were tested for 1400 seconds (14 100-sec. trials). For the first two trials (200 sec.) OE set prevailed; for the next 12 trials (1200 sec.) *S* shifted between RCW and TD sets: trials 3, 5, 7, 9, 11, 13 were RCW set; trials 4, 6, 8, 10, 12, 14 were TD set. Each *S* gave 280 reports for a total of 1400. For each *S* significantly more reports

of *rcw.* were given during the RCW set, compared to TD set, while the reverse was true for *td.* reports. The results are presented as the mean cumulative frequency of reports for the group during the 1400-sec. period. See Figure 1.



The straight line is the maximum possible rate. The horizontal axis refers to the serial position of the report during the experiment, i.e., first reports, second, etc. At the top of the graph the occurrence of the various instructional sets is indicated.

D. SUMMARY

In six experiments using various designs and various instructional sets *Ss* reported a given perceived motion more often when instructed to perceive that motion. The effects of set were evident during continuing sets (5-10 min. duration) as well as during different sets alternately activated for shorter periods (75-100 secs.) The reproducibility of these effects among many kinds of subjects indicated that the influence of set is of a general nature. These results extend the known effects of instruction to these perceived motion phe-

nomena, pointing up the necessity of controlling instructions and suggestions when studying these phenomena as well as indicating that, analogous to motor acts, perceptual Gestalten elicited by a constantly rotating stimulus can be controlled by the perceiving subject.

REFERENCES

1. AMES, A. Visual perception and the rotating trapezoidal window. *Psychol. Monog.*, 1951, **65** No. 324, 3-31.
2. MILES, W. R. Movement interpretation of the silhouette of a revolving fan. *Amer. J. Psychol.*, 1931, **43**, 392-405.
3. MULHOLLAND, T. Motion perceived while viewing rotating stimulus objects. *Amer. J. Psychol.*, 1956, **69**, 96-99.
4. ———. The swinging disk illusion. *Amer. J. Psychol.*, 1958, **71**, 375-382.
5. PASTORE, N. Some remarks on the Ames oscillatory effect. *Psychol. Rev.*, 1952, **59**, 319-323.
6. WALLACH, H., & O'CONNELL, D. N. The kinetic depth effect. *J. Exper. Psychol.*, 1953, **45**, 205-217.

Perception Laboratory
Veterans Administration Hospital
200 Springs Road
Bedford, Massachusetts

A BEHAVIORISTIC RATIONALE FOR THE DEVELOPMENT OF PSYCHOSOMATIC PHENOMENA*¹

Department of Psychology, Wayne State University

SHELDON J. LACHMAN

A. INTRODUCTION

Historically, the ideas of Cannon expressed in his emergency reaction notion (5, 6), of Selye in his formulation of the general adaptation syndrome (15, 16), and of Lindsley in his activation theory (11), have been exciting and significant conceptualizations in the psychology of emotion. Each conceptualization has paid special attention to the extensive physiological changes in emotional behavior and each has much influenced contemporary thinking about emotional behavior.

Writings of Dunbar (7, 8), of Alexander (1, 2), and of Grinker (10), among others, have effectively extended thinking about emotional behavior into the realm of psychosomatic phenomena, especially psychosomatic disorders.

In recent years, also, there has been extensive research work on psychosomatic phenomena, much of it clinical but increasingly more of it of an experimental nature, by such investigators as S. Wolf and H. G. Wolff (18), White, Cobb, and Jones (17), L. J. Saul (14), Binger and associates (3), Brady (4), and many others.

With so much in conceptual and empirical background, it may now be appropriate to integrate contributions from these various sources into a preliminary rationale concerning the development of psychosomatic phenomena. Such a rationale would have the advantages of (a) organizing and *explicitly identifying a set of variables* which apparently are of major significance in the field of psychosomatics; (b) providing a means for *encompassing in a coherent way much of what is now known about the area and means for thinking about established findings*; and (c) *providing the foundation for more systematically investigating the area and for further elaborating*

* Received in the Editorial Office on May 21, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ In a slightly different form, this paper was presented at a meeting of the Medical Sciences Section of the Michigan Academy of Science, Arts, and Letters under the title of "A Rationale for the Development of Psychosomatic Disorders."

This is Report No. 2 from the Comparative and Physiological Psychology Laboratory at Wayne State University.

The author is grateful to Wilson McTeer for his critical reading of the paper.

knowledge concerning the genesis and development of psychosomatic phenomena.

The rationale proposed in the present paper consists of a series of related propositions, supported to a greater or lesser degree² by empirical evidence. This rationale is not necessarily in conflict with the well-established theoretical views. In fact, several apparently different theories concerning the genesis of psychosomatic illness may find in the present article much to which they can readily subscribe, i.e., it is suggested that at least some of the propositions specified will represent processes about which there is converging agreement in the different theories. What is endeavored in this presentation is to express overtly and objectively a set of harmonious propositions concerning the origin and development of psychosomatic phenomena, to relate these propositions to each other and present them in a particular organized way, and to provide thereby a propositional structure from which new hypotheses are readily derivable and which can be checked and further developed through methodical investigation.

Our theory is stated strictly in behavioristic terms; it emphasizes the role of learning in the development of psychosomatic aberrations without minimizing the role of genetic factors or of nongenetic predisposing factors.

The essence of the theory proposed is that: *Psychosomatic manifestations result from frequent and/or prolonged and/or intense implicit reactions² elicited via stimulation of receptors.*

B. DEFINITION OF EMOTION AND VARIETIES OF PSYCHOSOMATIC PHENOMENA

A definition of emotion, or rather of emotional behavior, seems desirable for the purposes of this paper. *Emotional behavior refers to extensive and pronounced patterns of reaction in structures innervated by the autonomic nervous system in response to stimulation of receptors.* This definition indicates that (a) the reaction is typically not narrowly localized; (b) it involves implicit structures (the effects of which may not be immediately explicit); and (c) it is aroused originally by receptor stimulation and is not a consequence of exercise, bacterial invasion, tissue injury, internal pathology, drug injection, or other such conditions.

Extensive implicit reactions to sensory stimulation, then, is the specifiable

² The phrase "implicit reactions" refers to physiological reactions such as changes in heart rate, changes in blood pressure, changes in rate of stomach or duodenal movement, changes in gastric gland or adrenal gland secretion, etc. These are covert responses—responses which ordinarily are not directly apparent, but which may be detected by special procedures or instrumentation.

essence of the definition of emotional behavior; usually these physiological changes are in the direction of facilitating or preparing the organism for defense, that is, they are adaptive. If the physiological changes are of more than mild intensity and of more than brief duration, however, there may be further consequences relevant to the health of the organism. These are called *psychosomatic reactions*. At least three categories of psychosomatic phenomena can be distinguished:

1. *Constructive Psychosomatic Reactions*

These are physiological reactions to emotional stimulation which are of *special value* to the organism in *combatting illness or disease processes*; they serve to *counteract* the effects of illness symptoms or disease or are otherwise adaptive for the organism. Several examples may be specified: (a) reduction of internal hemorrhaging due to patterns of visceral vasoconstriction (and peripheral vasodilation) in response to emotional stimulation; (b) increased cardiac activity in the individual with serious bradycardic symptoms; (c) increased blood pressure in the patient with very low blood pressure; (d) decreased digestive activity as part of an emotional reaction in the case of the individual who has inadvertently swallowed a poisonous substance and then recognized what he has done.

2. *Destructive (or Pathological) Psychosomatic Reactions*

These are physiological changes induced by repeated or persistent emotional stimulation which (because of intensity level) are disadvantageous to the organism and are of sufficient chronicity or duration to be considered aberrations or dysfunctions. In other words, these reactions, though they may sometimes be completely or partially reversible, are so unsuitable or unfavorable to the individual that they are regarded as significantly deviant or abnormal. To put it still differently, this category includes pathological functional alterations and would be exemplified by certain varieties of hypertension, colitis, and asthmatic attacks.

3. *Psychosomatic Diseases*

These are actual organic or physical pathologies resulting from sustained, repeated, or intense reactions to emotional stimulation. In other words, the psychosomatic diseases are conditions in which there is relatively permanent structural change of a maladaptive kind, i.e., tissue damage and organic destruction. Duodenal ulcer resulting from physiological reactions to emo-

tional stimulation and permanent cardiac dilation to similar stimulation would be considered examples of psychosomatic diseases.

The generic term, psychosomatic phenomena, is used to refer to emotional reactions and to the three additional categories suggested. The terms psychosomatic disturbance or psychosomatic aberration or psychosomatic disorder may be used interchangeably to refer to both destructive psychosomatic reactions and to psychosomatic diseases.

C. OBJECTIVE RATIONALE FOR PSYCHOSOMATIC DISORDERS

The position to be developed which may be called an Autonomic Reaction Learning Theory of Psychosomatic Phenomena perhaps will be clarified if a few of the assumptions regarding emotional behavior are made explicit:

1. Regardless of the original or native stimuli which elicit emotional behavior, new cues may become conditioned or otherwise learned to serve as effective elicitors of intense internal behavior.

2. Intense implicit reactions to specific stimulus patterns, once aroused, tend to be extremely persistent; that is, recurrence of the specific eliciting stimulus patterns tends to produce particular implicit reactions repeatedly with little tendency toward extinction.

3. Overt accompaniments of emotional behavior are more readily "controlled" and directed by the organism than are the internal reactions.

The central thesis of our position is: Emotional behavior (intense internal reactions which initially in the history of the organism are aroused via receptor stimulation) may produce more or less permanent structural changes or may lead to conditions of chronic physiological activation—psychosomatic manifestations.

This position is further specified and elaborated in the following outline which attempts to consider the necessary reaction bases and the necessary stimulus bases for psychosomatic phenomena.

1. *The Necessary Reaction Bases for Psychosomatic Phenomena*

The anatomical-physiological mechanism necessary for the development of psychosomatic phenomena in an individual must possess at least two characteristics:

- a. Reaction variation.* One of these is the capacity of particular implicit response mechanisms to vary their activity. This may be variation in rate and/or in amplitude of response from a typical ("steady" or "constant" or "homeostatic") reactivity level.

b. Response predilection to explicit stimuli. The second of these is that response mechanisms which play a role in psychosomatic phenomena have the capacity to respond to explicit stimuli, either natively, or as a function of an associative process.

2. *The Necessary Stimulus Bases for Psychosomatic Phenomena*

a. Individual differences in threshold. The threshold of stimulation initially capable of provoking emotional reactions varies from one individual to another. It is possible that the range of native elicitors of emotional reaction also varies from one individual to another.

b. Emotional-stimulus learning.

(1). *The role of stimulus substitution learning.* On an associative basis, new effective stimuli may be developed. In other words, stimuli originally incapable of eliciting a response but which are closely associated temporally with effective emotion-provoking stimuli, may themselves sooner or later, come to be effective in producing emotional reactions.

(2). *The role of stimulus generalization.* An internal response which has been associated with a particular stimulus may come to be elicited by a variety of somewhat similar stimuli, i.e., stimuli somewhat like the originally learned stimulus, but varying along one continuum or another. Such stimuli become capable of eliciting an internal response (emotional reaction or aspect thereof) which they did not elicit prior to learning.

(3). *The role of symbolic stimuli.* Stimuli which in the personal history of the individual represent effective emotion-provoking stimuli, may themselves become emotion-provoking. At the human level, such symbolic stimuli are frequently, but not always, language stimuli.

(4). *The role of ideation.* Organismic effects of stimuli, probably largely in terms of central nervous manifestations, may persist in the form of "central percepts" or "ideas" which can be revived in the absence of the originally relevant external stimulus-situation, to produce a characteristic implicit emotional-reaction constellation.

c. The vicious circle. "Vicious Circle" phenomena may also occur. Once initiated, a psychosomatic event may produce stimuli which lead to implicit reactions that serve to intensify the psychosomatic event, which may elicit new stimuli producing further implicit reactions and further intensification of the psychosomatic event, etc.

For example, the noxious stimulation from a gastric ulcer may elicit implicit reactions including facilitated stomach acid secretion which intensifies that ulcerous condition thereby producing further noxious stimulation which

leads to further emotional reaction, and further irritation of the ulcer. Theoretically, and perhaps in fact, an individual may actually "worry" himself to death in such a "vicious circle" fashion.

D. GENERAL FACTORS IN THE GENESIS AND DEVELOPMENT OF PSYCHOSOMATIC PHENOMENA

In addition to the reaction bases and the stimulus bases outlined, it seems necessary to specify some principles regarding predisposing and precipitating determiners of psychosomatic phenomena.

1. *Predisposing Factors of Psychosomatic Phenomena*

a. Predisposing genetic factors. Physical structures of the organism are in the last analysis biologically transmitted (i.e., genetically inherited) structural characteristics of the organism. These include such characteristics as quality of heart muscle, of gastrointestinal tract, of tracheal apparatus, of endocrine glands, of blood vessels, etc. There may be many gradations of "strength" or durability of these structures as a function of genetic influence. Certain organic structures have a "low threshold" for or high susceptibility to deleterious structural modifications.

Every structure has an "inherent level of resistance" to the fields of forces which it encounters; there are vast individual differences. Some structures "break down" or "deteriorate" or "change" readily; others display greater resistance to change.

Differences in genetic predisposition are both interindividual and inter-systemic.

b. Predisposing environmental factors. Acquired physical modifications are consequences of direct physical interaction of organismic structures with the environment, and include the effects of such phenomena as physical assault, parasitic infection, neoplastic growths, foreign toxic agents, nutritional deficits, etc. These may alter inherited structure deleteriously or perhaps in some cases, raise rather than lower the "resistance level" of these structures.

The "level of resistance," i.e., threshold of injury, of an organic structure may be altered as a consequence of direct and indirect physical interaction with the environment.

2. *Precipitating Factors of Psychosomatic Phenomena*

The precipitation of a phenomenon which is to be classified as a psychosomatic disorder, according to our rationale, must be provoked by internal

behavior which in turn has been elicited by ordinary stimulation of receptor structures. These internal behavior patterns, i.e., emotional reaction patterns, may vary in extent, in frequency, in intensity, and in duration in producing such a phenomenon.

a. The extensity of the emotional reaction pattern. There are wide individual differences in the number of structures involved in an emotional reaction. Reacting structures may be few or many in number. There is variation from individual to individual and from time to time in the same individual.

The structures involved in emotional reaction patterns are the structures which are primarily involved in the development of psychosomatic disorders.

b. The frequency of the emotional reaction pattern. There are wide individual differences in the frequency of occurrence of emotional reactions.

The more frequently a given structure is involved in emotional reactions, the greater the likelihood of its being involved in a psychosomatic disorder, other things being equal.

c. The intensity of the emotional reaction pattern. There are wide individual differences in the "severity" of the change in the functioning of a structure during an emotional reaction. The greater the deviation or variation in functioning of a structure from its "normal" resting or homeostatic level, the greater the likelihood of its being involved in a psychosomatic disorder, other things being equal.

d. The duration of the emotional reaction pattern. There are wide individual differences in the duration of the emotional reaction pattern, although such reactions are typically of relatively short duration.

The longer a given structure is involved in an on-going emotional reaction pattern, the greater the likelihood of its being involved in a psychosomatic disorder, other things being equal. (This implies that the factors specified as precipitating factors may also operate as predisposing influences.)

E. THE GENESIS AND DEVELOPMENT OF PSYCHOSOMATIC PATHOLOGY

The basic notion of a psychosomatic disorder is that: Physical injuries or pronounced physiological dysfunctions may be elicited independently of externally induced physical means. Such an injury or dysfunction is a result of internal reactions elicited, at least initially, by means of sense organ stimulation.

In order for emotional reactions to assume pathological significance, such reactions must be intense and/or chronic. Which structure will ultimately be affected depends on (a) genetic factors; (b) predisposing environmental

influences; and (c) the particular structures involved in emotional reactions, and the magnitude (intensity, frequency, and duration) of their involvement.

F. A FEW DETAILS CONCERNING THE THEORY

What *structure* ultimately becomes involved in a psychosomatic reaction, then, depends on the biological condition of the structure (whether determined by innate or environmental factors), on the initial reactivity threshold of the organ, and on such learning factors as the development of stimulus-substitution reactions, generalization, symbolic stimuli, and ideation. The *magnitude* of the psychosomatic phenomenon is dependent on how frequently the structure is affected, how intensely it is affected, and how long it is affected.

All internal mechanisms are not activated to the same degree. Under various circumstances one group of structures, let us say peripheral blood vessels, or the heart, or stomach muscles, may be intensely activated while at other times different structures are maximally activated. As a consequence of such activation having occurred once, the immediate stimulus situation and other stimuli associated with it temporally are likely to be able to elicit similar reactions in the future. With frequent occurrence of the sequence, there is increasing likelihood of recurrence of the response to the stimulus-situation and an S-R pattern becomes firmly established. A small part of the total more complex stimulus-situation may come to elicit the reaction—redintegration. Temporally and spatially associated stimuli may also do so—stimulus substitution. Similar stimuli may do so—generalization. Symbolic stimuli—not similar in stimulus dimensions but similar in terms of meaning—may do so. Equivalent internal, i.e., representative or ideational, central nervous system effects, may do so. Essentially, these are learning phenomena.

G. SUMMARY STATEMENT ON THE AUTONOMIC REACTION LEARNING THEORY OF PSYCHOSOMATIC PHENOMENA

No single autonomic reaction pattern is emotional behavior; there are many possible autonomic patterns. With reference to the organs involved, these differences in pattern depend on different eliciting circumstances as well as different genetic and nongenetic predisposing variables and there are not merely intraindividual but there are also obviously interindividual differences in reaction pattern. New stimuli, through learning, may elicit emotional reaction patterns.

Much evidence at this time suggests the nature of the mechanisms involved in the development of psychosomatic phenomena beginning with receptors

and afferent structures and terminating with peripheral aspects of the autonomic nervous system, including also the structures influenced directly by the autonomic nervous system. To elucidate these mechanisms is a major objective of investigators.

Pertinent to this objective, specific knowledge which will be of value in understanding neural mechanisms in psychosomatic phenomena has emerged significantly from the work of Magoun (13), French and his associates (9), and MacLean (12), as well as from other published research.

A brief formalized framework such as the one herein proposed (*a*) can provide a means for encompassing in a general way much of what is now known; and (*b*) can provide the foundation for more systematic investigation of variables in the future.

Perhaps emotional reactions are basically constructive—defensive and adaptive. However, if emotional activity or reactivity is sufficiently frequent or prolonged or intense, it becomes maladaptive or destructive, leading to physiological aberration or structural damage to the organism, and even to death. Thus, an organism may be injured or destroyed by its own defenses. The propositions outlined provide the framework of a rationale suggesting how this might occur psychologically and biologically.

Theoretically, any bodily structure or function can become the end-focus of psychosomatic phenomena—but especially those *directly* innervated and regulated by the autonomic nervous system.

The values of particular frustrations, psychological trauma, conflicts, and other stresses in terms of emotional reactivity provoked, must be individually assessed for each person—but first more general research is necessary to establish more specifically the major parameters of influence.

REFERENCES

1. ALEXANDER, F. Fundamental concepts of psychosomatic research: Psychogenesis, conversion, specificity. *Psychosomat. Med.*, 1943, 5, 205-210.
2. ———. *Psychosomatic Medicine: Its Principles and Applications*. New York: Norton, 1950.
3. BINGER, C. A., ACKERMAN, N. W., COHN, A. E., SCHROEDER, H. A., & STEELE, J. M. Personality in arterial hypertension. *Psychosomat. Med. Monog.*, 1945, No. 8.
4. BRADY, J. V. Ulcers in "executive" monkeys. *Sci. Amer.*, 1958, 199 (3), 95-104.
5. CANNON, W. B. *Bodily Changes in Pain, Hunger, Fear, and Rage: An Account of Recent Researches into the Function of Emotional Excitement*. New York: Appleton, 1929.
6. ———. *The Wisdom of the Body*. New York: Norton, 1939.
7. DUNBAR, H. F. *Psychosomatic Diagnosis*. New York: Hoeber, 1943.
8. ———. *Emotions and Bodily Change* (3rd ed.) New York: Columbia Univ. Press, 1946.

9. FRENCH, J. D., PORTER, R. W., CAVANAUGH, E. B., & LONGMIRE, R. L. Experimental observations on "psychosomatic" mechanisms: I. Gastrointestinal disturbances. *A.M.A. Arch. Neurol. Psychiat.*, 1954, **72**, 267-281.
10. GRINKER, R. R. *Psychosomatic Research*. New York: Norton, 1953.
11. LINDSLEY, D. B. Emotion. In S. S. Stevens (Ed.), *Handbook of Environmental Psychology*. New York: Wiley, 1961.
12. MACLEAN, P. D. Studies on the limbic system ("visceral brain") and their bearing on psychosomatic problems. In Wittkower, E. D., & Cleghorn, R. (Eds.), *Recent Developments in Psychosomatic Medicine*. Philadelphia: Lippincott, 1954.
13. MAGOUN, H. W. An ascending reticular activating system in the brain stem. *A.M.A. Arch. Neurol. Psychiat.*, 1952, **67**, 145-154.
14. SAUL, L. J. Hostility in cases of essential hypertension. *Psychosomat. Med.*, 1939, **1**, 153-216.
15. SELYE, H. The general adaptation syndrome and the diseases of adaptation. *J. Clin. Endocr.*, 1946, **6**, 117-128.
16. ———. *The Stress of Life*. New York: McGraw-Hill, 1956.
17. WHITE, B. V., COBB, S., & JONES, C. M. Mucous colitis: A psychological and medical study of sixty cases. *Psychosomat. Med. Monog.*, 1939, No. 1.
18. WOLF, S., & WOLFF, H. G. Evidence on the genesis of peptic ulcer in man. In S. S. Tomkins (Ed.), *Contemporary Psychopathology*. Cambridge: Harvard Univ. Press, 1943.

Department of Psychology
Wayne State University
Detroit 2, Michigan

\$20.00 per annum
\$15.00 per volume
\$7.50 per half volume

QUARTERLY
Two volumes per year
Immediate publication

October, 1963
Volume 56, Second Half

(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a liberal contract, and may not be resold for a period of five years from the date of publication.

OCTOBER, 1963

(Manuscripts are printed in the order of final acceptance)

- Anxiety and achievement of intellectually gifted and creatively gifted children 251
By IRWIN FLESCHER
- Has the status of music changed in 30 years? 269
By PAUL R. FARNSWORTH
- The effect of psychopathology in children's drawing of the human figure 273
By THERON ALEXANDER
- The retention of sensory experience: I. Intensity 283
By H. E. KING
- The retention of sensory experience: II. Frequency 291
By H. E. KING
- The retention of sensory experience: III. Duration 299
By H. E. KING

(OVER)

Copyright, 1963, by The Journal Press
Provincetown, Massachusetts, U. S. A.

All rights reserved

Entered as second-class matter January 26, 1937, at the post-office at
Provincetown, Mass., under the Act of March 3, 1879
Second-class postage paid at Provincetown, Mass.

| | |
|---|-----|
| A preventive mental health program in the school setting: Description and evaluations | 307 |
| BY EMORY L. COWEN, LOUIS D. IZZO, HAROLD MILES, EARL F. TELSCHOW, MARY ANN TROST, AND MELVIN ZAX | |
| Clinical judgment | 357 |
| BY DAVID J. KING AND RICHARD F. MANEGOLD | |
| Lysergic acid diethylamide (LSD-25): XXXIV. Comparison with effect of psilocybin on the Siamese fighting fish | 363 |
| BY H. A. ABRAMSON, H. H. GETTNER, A. ROLO, AND G. DEAN | |
| Emotional behavior of the white rat following full-thickness thermal burns | 375 |
| BY FRED T. CALDWELL, JR., BETTYE M. CALDWELL, AND STANLEY E. MOLDOVAN | |
| Toward an automated technique of probing into emotional blocks | 385 |
| BY ALBIN R. GILBERT | |
| Self-disclosure and expressed self-esteem, social distance and areas of the self revealed | 405 |
| BY MAUREEN P. FITZGERALD | |
| The effect of observer redundancy on display monitoring equipment | 413 |
| BY DONALD W. BROWN | |
| Response set in objective achievement tests | 421 |
| BY ALICE GUSTAV | |
| Probability preferences and expected values | 429 |
| BY ALVIN SCODEL | |
| A study of the possible effect of differential light adaptation on measures of brightness enhancement | 435 |
| BY CHARLES M. BOURASSA AND S. HOWARD BARTLEY | |
| Anxiety, extraversion, and personality idiosyncrasy in delinquency | 441 |
| BY GEORGE R. PIERSON AND ROBERT F. KELLY | |
| The effect of group verbal suggestion and age on the perception of the Ames trapezoidal illusion | 447 |
| BY MOTHER JANE MARIE MCGEE | |
| Studies in the neurophysiology of learning: VII. Bar-pressing rate for cranial self-stimulation as related to frequency of rectangular pulses when Q is held constant | 455 |
| BY J. A. GENDERELLI, RONALD D. PRIDDY, AND JAMES R. AVERILL | |
| Sex differences among schizophrenics in the interpretation of the human diad | 473 |
| BY ROBERT H. KNAPP AND LINDA A. EHLINGER | |
| Transactional aspects of ego therapy | 479 |
| BY STEWART B. SHAPIRO | |

ANXIETY AND ACHIEVEMENT OF INTELLECTUALLY GIFTED AND CREATIVELY GIFTED CHILDREN*¹

East Williston Public Schools, New York

IRWIN FLESCHER

A. INTRODUCTION

An immediate concern with regard to recent explorations of the creative process is the limited generalizations that can be made from the findings. The discovery (2) that creativity is an important determinant of academic achievement, although there is a relatively low correlation between creativity and intelligence, appears destined to have far-reaching educational implications. In view of this, the atypical nature of the selected sample in the aforementioned study necessitates scientific caution in interpreting the significance of those results. The demonstration of equal superiority in scholastic achievement between a highly intelligent and a highly creative group was based on the failure to reject the null hypothesis. It is noteworthy that although the mean *IQ* scores were disparate, both groups represented an exceptionally high order of intellectual potential. There is a question as to whether differential achievement expectations were realistic in that situation, particularly since standardized achievement tests are deficient in measuring gifted vertical and horizontal development in areas beyond the normally prescribed curriculum. In a series of replicated procedures (15) with a variety of school populations, it was found that the Getzels-Jackson phenomenon is likely to occur under certain conditions but not under others. However, the reasons for the failure to reproduce the results in certain samples are not clearly evident.

The intent of the present study was to pursue the validity of implications concerning the comparative influence of unusual creative thinking and exceptional intellectual endowment on the learning process. It was felt that more meaningful observations and increased confidence in the findings would derive from research involving an adequate control group characterized by non-extraordinary intelligence and creativity. Furthermore, these indices

* Received in the Editorial Office on May 20, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ This research was supported by the East Williston Board of Education, New York. The author wishes to thank Dr. Neil V. Sullivan, Superintendent of Schools, for his interest and approval of the study.

would need to be respectively equated with non-talented intelligence as found in a creatively-gifted group, and non-talented creativity as found in an intellectually-gifted group. Scientific inquiry warrants inclusion of still another group in which subjects are simultaneously gifted in intelligence and creativity. It was stipulated in the previously cited studies (2, 15) that individuals outstanding in both respects were eliminated from the design. This omission assumes critical importance in the light of the theoretical proposition (6) that general intelligence and creativity may be respectively conceptualized as convergent- and divergent-thinking processes which represent two apparently antithetic factors of intellect. How formalized intelligence and creative ability interact to affect school achievement is a necessary dimension to examine.

Explanations regarding complex cognitive functioning inevitably lead to the search for intervening variables. The Getzels and Jackson study revealed that achievement motivation was not a distinguishing factor. In seeking clarification of the intelligence-creativity-achievement relationship, the present investigation was designed to determine to what extent the personality variable of anxiety is a mediating influence.

The degree to which operational anxiety may be a hidden selective factor when groups are constituted by levels of intelligence and creativity is of crucial importance in contrasting achievement measures. The interfering and the facilitating effects of differential anxiety on academic achievement have been established (12). There is reason to suppose that the effects of inordinate anxiety are no less disruptive and restrictive in relation to the elusive and fragile gift of productive creativity. Evidence of an inverse relationship between the personality trait of attitudinal rigidity and verbal creativity (1) support this position. The converse proposition logically follows that a low anxiety level is a necessary correlate of unhampered and spontaneous psychological processes characteristic of creative endeavor. This assumption is in keeping with the neo-psychoanalytic formulation (9) that creativity depends on an inner freedom from distortions and obstructions of neurotic origin. The dependence of divergent-thinking operations on personality dynamics is further suggested by the finding that ". . . creative persons are especially disposed to admit complexity and even disorder into their perceptions without being made anxious by the resulting chaos" (11, p. 489).

B. METHOD

1. *Measures*

In the early part of 1962 a variety of creative aptitude tests was administered to the sixth-grade population of the East Williston School District.

These divergent-thinking tasks were modeled after the theoretical constructs developed by Guilford (4, 5). The creativity battery included five separate tests; two of these were scored twice, for independent criteria of creative thinking. The seven subscores were converted to standard scores, and a composite measure was computed which yielded an index of Total Creativity.

A brief description of the divergent-thinking tasks and scoring procedures is presented below.

A word association test was given in which the children were required to respond to each of a list of 20 stimulus words. The test items were selected from a compilation of the 570 most commonly used words in the English language (10). For every stimulus, after discounting irrelevant and bizarre answers, a statistical count was made of the frequency of mention of each association. The only responses credited were words or phrases that were one of a kind. The score for each child gave a measure of Remote Associations.

As a measure of nonverbal creativity the students were instructed to "draw an old man." Inspection of the projective material revealed a consensus in the modes of expressing this concept, e.g., a high frequency of beards, canes, and eyeglasses. Productions which digressed from the standard perspective, and which were judged to be novel or clever, were given appropriate credit as Divergent Drawings.

A test requiring the designation of different uses for each of six familiar objects was administered. Common and unusual uses were determined by statistical frequency. Two scores were derived from this test. The crediting of unusual evocations yielded a score for Unique Uses. In addition, the test replies, exclusive of the unusual answers, were re-evaluated. The enumerated common responses to each stimulus were inspected to ascertain for each child the number of changes in *set* from one type of content category to another. This supplied a measure of Spontaneous Flexibility.

A task of devising diversified titles for a given story plot was included in the battery. Rare and clever titles provided a score for Original Titles. In addition, after discounting original responses, the number of obvious or ordinary titles offered determined the degree of Ideational Fluency.

The final test in the creativity series was a written composition. The children were requested to project themselves into the future and write an account about themselves as ladies or gentlemen 25 years hence. The entire set of future autobiographies was reread many times to establish stereotyped content and the incidence of repetitive themes. Uncommon and original ideas, and novel reorganization of familiar experiences, were the criteria used in identifying Imaginative Compositions.

Subsequent to the presentation of the divergent-thinking tests, the pupils were required to respond to two anxiety inventories (12). The General Anxiety Scale for Children (GASC) and the Test Anxiety Scale for Children (TASC) were administered in that sequential order. Scoring was in accordance with the standardized procedures, and the resultant ratings gave evidence of the extent of General Anxiety and Test Anxiety for each child. The two scales were also combined into a Total Anxiety score.

Aside from the obtained information on creativity and anxiety, the current results of the routine school testing program were made available. These included data on intelligence and scholastic achievement.

The California Test of Mental Maturity (CTMM) supplied the measure of intellectual potential (*IQ*). It is noteworthy that the mean intelligence level of the available student population fell within the upper segment of the above-average range, and therefore, the incidence of high *IQ* scores was comparatively greater in this sample than is found in a nationally average population.

Achievement indices were derived from the results of the Metropolitan Achievement Tests (MAT). The Intermediate Battery yielded standard scores for Word Knowledge, Reading, Spelling, Language, Language Study Skills, Arithmetic Computation, Arithmetic Problem Solving and Concepts, Social Studies Information, Social Studies Study Skills, and Science. A composite score for Total Achievement, based on the ten subscores, was also computed.

2. Groups

The assembled data on *IQ* and Total Creativity were the indices used in the selection of four subsamples from the tested population. The stringent selection criteria precluded the formation of adequate subsamples by sex. Boys and girls were approximately equal in number in each group, thereby minimizing sex differences as a source of variation.

The groups were constituted as follows:

Ss characterized by high intelligence as defined by an *IQ* of 130 or above, but who were not in the top quartile in creativity, comprised the *Intellectually-Talented* ($N = 28$). Conversely, *Ss* who scored in the top quartile on tests of creative thinking, but whose *IQ* scores fell below 130, represented the *Creatively-Talented* ($N = 29$). A third group of *Ss*, identified by *IQ* scores of 130 or higher, and by creativity scores ranging within the top quartile, was designated the *Twice-Talented* ($N = 24$).

The procedure in the formation of the fourth group included additional

criteria. This group represented *Ss* with *IQ* scores below 130 and creativity scores below the top quartile. After the selection of the three talented groups on the basis of exceptional intelligence and/or exceptional creativity, the entire remaining population sample fell into the non-talented category. Since the intent was the establishment of a representative control group in accordance with a balanced design, the array of non-gifted scores in the two singly-talented groups had to be accounted for. Hence, each *S* was selected from the remaining non-exceptional population whose *IQ* score exactly matched or closely approximated one *IQ* score in the group designated the *Creatively-Talented*, and who simultaneously possessed a creativity score which exactly matched or closely approximated the creativity measure of one individual in the group representing the *Intellectually-Talented*. The resultant group was the *Non-Talented* ($N = 29$), a subsample characterized by means and standard deviations for intelligence and creativity which were of the same magnitudes as their corresponding non-exceptional counterparts in the two otherwise gifted groups.

3. Analysis

In accordance with the plan of the investigation the four counterbalanced groups were combined into a two-way classification system. Table 1 illustrates the scheme of the 2×2 factorial design. Each of the two independent variables, intellectual endowment and creative thinking, is represented by two complementary aptitude levels (relatively defined as *high* and *low*). The arrangement provided for the simultaneous evaluation of differences between the row means and between the column means. Furthermore, the factorial approach permitted comparison of the two variables in all combinations, allowing the determination of whether there was an interaction between the row and column components, or whether each exercised an influence apart from the other. Thus, the design yielded estimates of the variance due to each of the main effects, *intelligence* and *creativity*, and to the *intelligence* \times *creativity* interaction. A series of two-way analyses of variance and their corresponding *F* ratios were computed for the various measures of anxiety and achievement.

The analysis of variance provided the method of assessing differences in anxiety and in achievement between groups of children classified by levels of intelligence and creativity. In order to pursue relevant implications of the present research, it was also important to investigate similarities and differences between various sets of tests for the combined study sample. This was accomplished by the technique of multiple-factor analysis.

TABLE 1
2 X 2 FACTORIAL DESIGN

| | | CREATIVITY | |
|--------------|------|------------|-----|
| | | high | low |
| INTELLIGENCE | high | IC | Ic |
| | low | iC | ic |

Note: The four subsamples are designated symbolically. In contrast to lower case letters, upper case letters denote a gifted degree of intelligence or creativity:

Ic represents the *Intellectually-Talented*;

iC represents the *Creatively-Talented*;

IC represents the *Twice-Talented*;

ic represents the *Non-Talented*.

In the factor-analytic method only such factors emerge as are entered in the battery. In the present design the various anxiety, intelligence, and creativity measures, by virtue of the assumption of their relationships with achievement, were alleged to be determinants of scholastic performance. Inclusion of the achievement tests provided the criterion (academic success). The major purpose of the factor analysis was to see which of the other tests in the battery were loaded with the same factors with which the criterion was heavily loaded. It was felt that multiple-factor analysis would lend additional meaning and clarification to the outcome of the two-way analysis of variance.

Accordingly, product-moment correlations were computed among 23 variables including all the available scores of anxiety, achievement, intelligence, and creativity. Prior to the derivation of the factor-analytic data the composite measures of Total Anxiety and Total Achievement were removed from the correlation matrix, since whatever factors they represented were already present in their corresponding subtests. The matrix of intercorrelations for the remaining 21 variables was factored by the centroid method, and the extracted factors subsequently rotated to simple structure by the varimax procedure.

C. RESULTS

Measures of central tendency and variability for the IQ and Total Creativity scores of the four research groups are shown in Table 2. A

preliminary analysis of variance and corresponding tests of significance for paired contrasts indicated that all "high" vs. "low" *IQ* contrasts differed beyond the .01 level, and that there were no differences between the two "high" *IQ* or between the two "low" *IQ* measures; the very same relationships were established when the Total Creativity means were similarly scrutinized. Thus, the comparative identities of the four select groups were statistically delineated.

TABLE 2
MEANS AND STANDARD DEVIATIONS OF *IQ* AND TOTAL CREATIVITY MEASURES

| Group | <i>N</i> | Intelligence | | Creativity | |
|-------|----------|--------------|-----------|------------|-----------|
| | | \bar{X} | <i>SD</i> | \bar{X} | <i>SD</i> |
| Ic | 28 | 140.46 | 7.70 | 5.89 | 2.35 |
| iC | 29 | 115.55 | 8.98 | 13.28 | 2.59 |
| IC | 24 | 142.08 | 8.71 | 13.58 | 2.15 |
| ic | 29 | 115.68 | 8.81 | 5.82 | 2.19 |

Note: All I vs. i and C vs. c contrasts are significant beyond the .01 level.

1. Analysis of Variance

Table 3 presents the results of the analyses of variance for the three anxiety indices. The findings were uniformly negative, since nonsignificant results were obtained for both main effects and their interaction. In no instance was the null hypothesis in danger of rejection. This indicates a lack of supportive evidence that either of the two levels of intelligence or of creativity, or any combination of these factors, had a differential effect on the degree of measured anxiety. It was established that any concomitant differences in achievement in this research may not be ascribed to the mediating influence of anxiety.

The analyses of variance for the 10 subscores of achievement and the total composite score are summarized in Table 4. The findings were consistent in all cases; the main effect of intelligence was highly significant beyond the .01 level of confidence, while the effects of the creativity factor and the interaction were nonsignificant. As anticipated, intelligence is portrayed as a powerful determinant of academic achievement. However, the data did not allow the inference to be drawn that creativity is significantly related to scholastic aptitude. The Getzels-Jackson phenomenon was not evidenced.

2. Intercorrelations

The intercorrelation coefficients for the 23 variables are arrayed in Table 5. Anxiety is seen as inversely related to achievement, but the low negative

TABLE 3
SUMMARY OF ANALYSES OF VARIANCE FOR ANXIETY DATA

| Variable | Source | <i>df</i> | <i>MS</i> | <i>F</i> |
|-----------|------------------------------|-----------|-----------|----------|
| Gen anx | Intelligence (<i>IQ</i>) | 1 | 10.17 | .24 |
| | Creativity (<i>Cr</i>) | 1 | 28.18 | .67 |
| | <i>IQ</i> \times <i>Cr</i> | 1 | 34.41 | .82 |
| | Within (<i>W</i>) | 106 | 42.07 | |
| Test anx | <i>IQ</i> | 1 | 19.95 | .74 |
| | <i>Cr</i> | 1 | 29.65 | 1.10 |
| | <i>IQ</i> \times <i>Cr</i> | 1 | .00 | .00 |
| | <i>W</i> | 106 | 27.01 | |
| Total anx | <i>IQ</i> | 1 | 58.65 | .54 |
| | <i>Cr</i> | 1 | 115.61 | 1.08 |
| | <i>IQ</i> \times <i>Cr</i> | 1 | 31.92 | .30 |
| | <i>W</i> | 106 | 107.48 | |

Note: None of the *F* ratios approaches significance.

correlations were only evidenced in a few of the achievement variables. The data suggest a slight tendency for manifest anxiety to intrude on academic achievement. Anxiety appeared for the most part to be unrelated to intelligence or creativity, with the one exception that Ideational Fluency was correlated with Test Anxiety. Apparently the presence of anxiety in examination situations is somewhat influential in stimulating the number of answers or ideas on tests where multiple responses are encouraged.

The intercorrelations among all 11 achievement categories were high positive, and all the MAT subscales were just as closely associated with intelligence. There were some meager positive relationships between various achievement measures and certain creativity subtests, notably Unique Uses and Imaginative Compositions. Of interest were the consistently low negative correlations between the achievement scales and Remote Associations.

IQ was also found to be inversely related to Remote Associations; however, the CTMM scale was discovered to be unrelated to the other creativity measures, with the further exception of a slight positive correlation with Imaginative Compositions.

The intracreativity correlations indicated moderate relationships between certain subscales, and no apparent relationships between others. The range of these subtest correlations (not including the composite Total Creativity variable) was from $-.19$ to $+.36$, with four *rs* significant at the .05 level and four *rs* significant at the .01 level. The presence of 13 nonsignificant *rs* implied that the creativity battery may have been measuring several un-

TABLE 4
SUMMARY OF ANALYSES OF VARIANCE FOR ACHIEVEMENT DATA

| Variable | Source | df | MS | F |
|-------------|----------------------------|-----|-----------|---------|
| Word know | Intelligence (<i>IQ</i>) | 1 | 2569.26 | 49.65* |
| | Creativity (<i>Cr</i>) | 1 | 82.41 | 1.59 |
| | <i>IQ</i> × <i>Cr</i> | 1 | 30.58 | .59 |
| | Within (<i>W</i>) | 106 | 51.75 | |
| Reading | <i>IQ</i> | 1 | 2337.05 | 41.42* |
| | <i>Cr</i> | 1 | 27.47 | .49 |
| | <i>IQ</i> × <i>Cr</i> | 1 | 18.59 | .33 |
| | <i>W</i> | 106 | 56.42 | |
| Spelling | <i>IQ</i> | 1 | 2669.05 | 37.61* |
| | <i>Cr</i> | 1 | .85 | .01 |
| | <i>IQ</i> × <i>Cr</i> | 1 | 3.12 | .04 |
| | <i>W</i> | 106 | 70.97 | |
| Language | <i>IQ</i> | 1 | 3899.48 | 59.25* |
| | <i>Cr</i> | 1 | 75.44 | 1.15 |
| | <i>IQ</i> × <i>Cr</i> | 1 | 25.14 | .38 |
| | <i>W</i> | 106 | 65.82 | |
| Language sk | <i>IQ</i> | 1 | 1919.14 | 28.45* |
| | <i>Cr</i> | 1 | 44.79 | .66 |
| | <i>IQ</i> × <i>Cr</i> | 1 | 7.16 | .10 |
| | <i>W</i> | 106 | 67.45 | |
| Arith comp | <i>IQ</i> | 1 | 5620.03 | 82.88* |
| | <i>Cr</i> | 1 | .11 | .00 |
| | <i>IQ</i> × <i>Cr</i> | 1 | 1.48 | .02 |
| | <i>W</i> | 106 | 67.81 | |
| Arith conc | <i>IQ</i> | 1 | 4310.33 | 85.00* |
| | <i>Cr</i> | 1 | 5.06 | .10 |
| | <i>IQ</i> × <i>Cr</i> | 1 | 16.43 | .32 |
| | <i>W</i> | 106 | 50.71 | |
| Soc stud | <i>IQ</i> | 1 | 1941.64 | 40.31* |
| | <i>Cr</i> | 1 | 26.10 | .54 |
| | <i>IQ</i> × <i>Cr</i> | 1 | 46.08 | .96 |
| | <i>W</i> | 106 | 48.17 | |
| Soc stud sk | <i>IQ</i> | 1 | 2780.13 | 37.72* |
| | <i>Cr</i> | 1 | 35.75 | .48 |
| | <i>IQ</i> × <i>Cr</i> | 1 | 49.01 | .66 |
| | <i>W</i> | 106 | 73.70 | |
| Science | <i>IQ</i> | 1 | 2258.11 | 51.28* |
| | <i>Cr</i> | 1 | 32.06 | .73 |
| | <i>IQ</i> × <i>Cr</i> | 1 | 118.65 | 2.69 |
| | <i>W</i> | 106 | 44.03 | |
| Total ach | <i>IQ</i> | 1 | 293622.32 | 100.61* |
| | <i>Cr</i> | 1 | 1493.66 | .51 |
| | <i>IQ</i> × <i>Cr</i> | 1 | 2190.03 | .75 |
| | <i>W</i> | 106 | 2918.55 | |

* Significant beyond the .01 level.

TABLE 5
23 X 23 INTERCORRELATION MATRIX
(N = 110)

| Variable | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
|----------------|----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|-----|-----|-----|-----|-----|
| 1 Gen anx | 57 | 91 | -17 | -06 | -09 | -03 | -06 | -08 | -08 | -21 | -22 | -23 | -15 | -11 | 06 | 04 | 05 | -01 | 04 | 16 | | |
| 2 Test anx | | 86 | -28 | -18 | -15 | -14 | -12 | -09 | -07 | -22 | -11 | -30 | -20 | -16 | 03 | 07 | 14 | 06 | -18 | 27 | -17 | 10 |
| 3 Total anx | | | -25 | -13 | -13 | -09 | -10 | -09 | -09 | -24 | -19 | -29 | -19 | -15 | 05 | 06 | 10 | 03 | -06 | 23 | -12 | 12 |
| 4 Word know | | | | 76 | 73 | 70 | 64 | 53 | 59 | 64 | 53 | 65 | 83 | 63 | -24 | -04 | 24 | 14 | 13 | 02 | 30 | 12 |
| 5 Reading | | | | | 66 | 68 | 68 | 49 | 54 | 63 | 58 | 64 | 82 | 61 | -24 | -04 | 22 | 09 | 19 | 05 | 20 | 12 |
| 6 Spelling | | | | | | 69 | 56 | 51 | 57 | 55 | 45 | 55 | 78 | 59 | -23 | -03 | 13 | 00 | 15 | -04 | 11 | 01 |
| 7 Language | | | | | | | 71 | 67 | 68 | 53 | 57 | 60 | 85 | 68 | -27 | 01 | 15 | 16 | 13 | 09 | 23 | 14 |
| 8 Language sk | | | | | | | | 57 | 65 | 55 | 56 | 66 | 81 | 56 | -28 | 03 | 19 | 13 | 07 | 07 | 26 | 15 |
| 9 Arith comp | | | | | | | | | 84 | 54 | 60 | 51 | 79 | 74 | -30 | -07 | 21 | -03 | 00 | 11 | 15 | 04 |
| 10 Arith conc | | | | | | | | | | 63 | 62 | 60 | 84 | 76 | -25 | 02 | 22 | 03 | 00 | 12 | 17 | 09 |
| 11 Soc stud | | | | | | | | | | | 63 | 62 | 60 | 76 | -17 | 02 | 12 | 00 | 02 | 08 | 24 | 08 |
| 12 Soc stud sk | | | | | | | | | | | | 60 | 77 | 57 | -19 | 08 | 20 | 07 | -05 | 02 | 28 | 07 |
| 13 Science | | | | | | | | | | | | | 76 | 59 | -19 | -02 | 09 | -05 | 06 | -15 | 20 | -07 |
| 14 Total ach | | | | | | | | | | | | | 79 | 62 | -24 | 00 | 22 | 06 | 08 | 05 | 26 | 09 |
| 15 IQ | | | | | | | | | | | | | | 79 | -30 | 01 | 16 | 02 | 16 | 02 | 21 | 04 |
| 16 Rem assoc | | | | | | | | | | | | | | | -28 | 18 | 06 | 05 | 15 | -05 | -19 | 37 |
| 17 Diver draw | | | | | | | | | | | | | | | | | 05 | 23 | -07 | -01 | 19 | 36 |
| 18 Unique uses | | | | | | | | | | | | | | | | | | 36 | 16 | 14 | 26 | 62 |
| 19 Spont flex | | | | | | | | | | | | | | | | | | | 01 | 35 | 25 | 65 |
| 20 Orig titles | | | | | | | | | | | | | | | | | | | | -15 | 18 | 35 |
| 21 Ideat flu | | | | | | | | | | | | | | | | | | | | | 19 | 43 |
| 22 Imag comp | | | | | | | | | | | | | | | | | | | | | | 41 |
| 23 Total cr | | | | | | | | | | | | | | | | | | | | | | |

Note: Decimal points have been omitted.

$r = \pm .19$ required for significance at the .05 level.

$r = \pm .25$ required for significance at the .01 level.

related elements, an observation which was subsequently confirmed by the factor-analytic data.

3. Factor Analysis

Six independent factors were isolated as a result of factor analyzing the intercorrelations among 21 scales. The rotated factor matrix is given in Table 6. The factors are presented in the traditional manner, listing each of the six in order of magnitude. Tests with loadings of $\pm .30$ or more were considered significant in naming a factor with confidence. Analysis of the loadings in each column resulted in the following factor identifications.

TABLE 6
ROTATED FACTOR LOADINGS

| Variable | I | II | III | IV | V | VI | R^2 |
|-------------|------|------|------|------|------|------|-------|
| Gen anx | -.08 | .02 | -.02 | .75 | -.07 | .08 | .59 |
| Test anx | -.19 | .16 | .00 | .64 | .09 | -.14 | .50 |
| Word know | .82 | .14 | .00 | -.20 | -.20 | .08 | .78 |
| Reading | .82 | .09 | -.01 | -.07 | -.20 | .11 | .74 |
| Spelling | .79 | -.04 | -.05 | .00 | -.16 | .11 | .68 |
| Language | .85 | .12 | .04 | .08 | -.03 | .00 | .75 |
| Language sk | .79 | .11 | .09 | -.02 | -.06 | -.05 | .65 |
| Arith comp | .76 | .09 | -.04 | .04 | .49 | -.15 | .85 |
| Arith conc | .80 | .09 | .03 | .02 | .42 | -.08 | .84 |
| Soc stud | .70 | .04 | .07 | -.26 | .09 | -.03 | .57 |
| Soc stud sk | .67 | .06 | .14 | -.21 | .20 | -.10 | .56 |
| Science | .76 | -.14 | .05 | -.25 | .02 | -.01 | .66 |
| <i>IQ</i> | .81 | .02 | .06 | -.03 | .25 | .02 | .72 |
| Rem assoc | -.34 | .05 | .06 | -.03 | .11 | .49 | .37 |
| Diver draw | -.03 | -.04 | .82 | .11 | .07 | .19 | .73 |
| Unique uses | .17 | .57 | .03 | .00 | .14 | .32 | .47 |
| Spont flex | .03 | .68 | .16 | .01 | -.12 | .07 | .51 |
| Orig titles | .12 | .08 | .02 | -.01 | -.10 | .46 | .24 |
| Ideat flu | .01 | .55 | -.04 | .20 | .05 | -.24 | .41 |
| Imag comp | .21 | .41 | .56 | -.24 | -.09 | -.15 | .62 |
| Total cr | .02 | .77 | .32 | .10 | .08 | .49 | .95 |

Note: Loadings of $\pm .30$ or greater are italicized.

Factor I is a substantially loaded factor. All ten achievement subtests had loadings of $+.67$ or higher, and this cluster represents the criterion of academic success. The high loading of $+.81$ for *IQ* reflects the validity of the synonymous reference to intelligence tests as measures of scholastic aptitude. Conversely, the achievement tests are seen as measuring a large component of formalized intelligence. This factor may therefore be appropriately described as learning ability or *general intelligence*. A moderate loading of $-.34$ for Remote Associations on Factor I is the only suggestion of an antithetic relationship between convergent thinking and an alleged

creativity variable. The loadings of the other creativity scales on this factor were negligible.

Factors II, III, and VI are specific factors relative to divergent thinking. Since all the noncreativity variables had insignificant loadings on these factors, the three columns were essentially descriptive of the factorial dimensions of the creativity battery. That several orthogonal factors were observed to contribute to the variance of the Total Creativity scores was already suspicioned from the nature of the creativity subtest intercorrelations. The three creativity factors do not lend themselves to easy interpretation, despite the fact that certain subtests of productive thinking were relatively pure measures of one factor, and entirely unrelated to the factors represented by the other two columns.

The chief loading on Factor II was Spontaneous Flexibility. Of the four variables measuring Factor II, Spontaneous Flexibility and Ideational Fluency were the scales not significantly weighted in the other two creativity columns of the matrix. This permits Factor II to be identified as a *flexibility-fluency* factor. The duality illustrates an interdependence between shifts in thought perspective and the generation of ideas. It is logical to suppose that performance on the tasks representing the other two scales loaded on this factor, evoking unusual uses and writing imaginative compositions, would be determined by the flexibility and fluency of verbal thought processes.

Factor III is somewhat paradoxical in content. Nonverbal and verbal creative productions were both loaded on the same factor. Divergent Drawings and Imaginative Compositions were freer, less structured tests than the rest of the creativity battery, and are representative of clinical projective techniques. The available recourse to inner fantasy activity through the media of imagery or ideas is the distinguishing feature of these tasks. It seems appropriate to refer to Factor III as *fantasy expression*.

Three test variables, Remote Associations, Original Titles, and Unique Uses, were weighted on Factor VI. The principal characteristic of this verbal factor centers about the novel response. Factor VI is therefore labeled *originality*.

The definition of Factor IV is relatively clear. Substantial loadings for the two anxiety scales define this factor as *anxiety*. While the analysis of variance negated the hypothesis that anxiety is an intervening influence, it is noteworthy that several achievement scales and the Imaginative Compositions test revealed slight, but probably insignificant, negative loadings on this factor.

Factor V is evidently a subset of achievement. The significant loadings on this factor were the two arithmetic scales. Their quantitative nature leads to the conclusion that this is a *numerical* factor.

D. DISCUSSION

The delicate complexity of the creative process seems vulnerable to the intrusion of personality dynamics. There is no serious dispute over the contention that nonintellective, personality factors are related to creative potential. However, the specific nature of this relationship is still relatively undefined. The present study explored the reasonable proposition that anxiety plays an intervening role in creative productivity. The results presented no convincing evidence that either general or test anxiety significantly affects performance on certain specified tasks of creativity.

A limiting feature of the present design may have been the use of objective inventories to measure anxiety. The GASC is concerned with diffuse emotionality, while the TASC items deal with physiological concomitants and habitual ways of behaving. The question arises whether an objective approach can adequately assess such an unconsciously determined, ego-involved variable as anxiety, particularly at a level of interaction with unstable creative energy. The depth of measurement of a personality trait is a function of the kind of test administered. It is conceivable that a more indirect, subjective-projective method might have provided a measure with increased sensitivity.

The hypothesis of a creativity-anxiety link is important enough to warrant further exploration. It is not sufficient to discover what level of anxiety is optimally related to creative productivity; an equally significant inquiry is the degree to which disruptive tensions and anxieties are generated by external pressures on the individual to constrict and suppress his creative strivings. The forces at work to redirect divergent processes into compatible, conforming behavior are powerful social motives. That many creative personalities are not readily understood or socially adaptable is suggested by an insightful study of creative scientists, (13) which described such characteristics as a repression of emotionality and a low level of interpersonal involvement in human relations. Torrance (16) has pointed out that creative children are likely to feel isolated and psychologically estranged from parents, teachers, and classmates. If this be true, then many home and school situations must give rise to serious inner conflicts and resultant anxieties in children with exceptional creative potential. A greater psychological understanding of the personality dimensions of creative children is one of the prerequisites for the reduction and prevention of their conflicting social needs and characteristic psychic turbulence. This knowledge is crucial to a rapidly evolving goal in contemporary education—the identification, preservation, and enhancement of creative talent.

Despite the present observation that anxiety was not significantly related to divergent thinking, the importance of creativity for personality-as-a-whole is undeniable; however, the importance of creativity for formal learning and school achievement, as we shall see, must be seriously questioned.

The consistent demonstration in the analysis of variance that the two differentiated levels of intellect yielded highly significant differences on all the achievement subtests is an indication that the academic needs of the intellectually-gifted students in this population sample were being met. It should be noted however, that the correlations between intelligence and all the achievement scales were substantial, and, that the factor-analytic data revealed a common loading for both variables. This is not surprising since the same kinds of problems and test items are used on both types of tests; furthermore, intelligence scales are often standardized and validated in terms of scholastic achievement. It is characteristic of both kinds of tests that they involve the logical ordering of many facts in order to "converge" upon a single, correct solution. The present findings provide empirical evidence that performance on standardized achievement tests involves convergent-thinking operations associated with the measurement of intelligence.

There is a sharp contrast between the highly significant main effect of intelligence and the decidedly nonsignificant F ratios regarding creativity. The results did not confirm the hypothesis that creative potential is a determinant of academic achievement. Inspection of the intercorrelation matrix provides additional information relevant to this finding. There was for the most part, a lack of positive relationships between the creativity measures and intelligence (one creativity scale was inversely related to IQ). Data supplied by Getzels and Jackson (3, p. 20) show moderate correlations, significant beyond the .01 level, between intelligence and a majority of the five creativity scales employed in their study. It is suggested that one reason for their resultant creativity-achievement relationship was the component of general intelligence selectively operating in their battery of creativity tests. The more "intelligence-free" characteristic of the creativity battery administered in the present study serves to partially explain the obtained negative results. The problem of whether and how creative aptitude influences the learning process bears further investigation.

A compelling question regarding research on creativity is why original thinking should be expected to be a relevant factor in scholastic achievement, the nature of the latter being so markedly convergent in content. Formal learning is very unlike original and creative activity, and convincing evidence of a significant relationship is still wanting. At present, intelligence tests are

referred to as tests of *scholastic aptitude*. On the basis of current knowledge, it is an unwarranted assumption to broaden the concept of scholastic aptitude to include creative potential. In fact, an inverse relationship between convergent and divergent processes is just as likely to be the case. It has been shown (7) that the individual who avoids improbable associations in preference for the highly probable is not likely to develop highly creative solutions to problems. In the present research, the significant negative loading for Remote Associations on the factor identified as *general intelligence* is further indication of a possible antagonistic relationship.

An important observation concerning the nature of the creativity battery was furnished by the factor analysis. The loadings emphasized the fact that some of the creativity scales were sole measures of one of the three isolated factors of divergent thinking—*flexibility-fluency*, *fantasy expression*, *originality*. The intercorrelations also revealed a lack of relationship between various creativity measures. Different subtests were totally independent measures of different things. Apparently, certain divergent-thinking tasks are also widely divergent from each other. The absence of common characteristics in the assortment of scales may serve to explain the inconsistent findings on creativity among different investigators, and is one plausible reason for the nonsignificant findings of the present endeavor. The problem has been clearly stated in a recent analysis of the meaning of creativity measures:

- If tests of different "creativity" sub-abilities show very low correlations, then which children will be identified as creative will depend very heavily upon the particular collection of subtests that is used. . . . That is, one group of "creative" children, identified by one set of tests, may score high on tests of academic achievement, whereas another, chosen by a different battery, may not (14, p. 423).

It is obvious that validity for many of the newly devised measures of creative potential is purely an assumption. The investigators of a recent study of divergent thinking concluded that they were "skeptical of research which indicates wide variability among students in creativity, when the variability is inferred solely from information gathered in testing situations of short duration" (8, p. 75). The questionable nature of the creativity indices demands restraint in the categorical labeling of individuals as more or less "creative." Any premature identification is likely to be misleading. There is the urgent reminder that the multiple measures of creativity are still in the experimental stage. It is unfortunate that, because of the high valuation on both intellectual pursuit and creative talent in our culture, doubtful inferences are being eagerly adopted as settled facts.

The nature of original thinking indicates that the new-found techniques of assessment are not sufficiently developed to adequately account for many elusive elements of creative endeavor. It is unlike convergent or rational thinking which conforms to a logical model and is more easily accessible to evaluation. Concerning creativity, what else can we currently uncover about the range and quality of hidden resources and potentialities of the student in the present study who responded that an automobile tire could be used as "a tarnished, oversized halo"? Thinking which diverges poses the riddle of the multiple-headed organism. The very concept of diversity suggests that there are testing innovations yet to come before application to the educative process may be meaningful and valid.

The detection and nurturance of productive originality is a laudable goal, but the practical value of creative development is not easily realized. It is not known, for example, what transfer value a high order of creative aptitude in childhood has for later life. In certain respects such a gift is not perceived as a particularly advantageous attribute. Getzels and Jackson (3) revealed that while parents and teachers ranked creativity among the first three qualities of gifted children, neither parents nor teachers included creativity among the first three qualities needed for success as adults; instead, they stressed personality attributes. The relevance for creativity is apparent however, since the factors of productive thinking are intimately linked to personality.

The emerging question as to how creative talent may be nurtured in the educational setting should be related to the purpose for which the productive-thinking abilities are being sought. If the specified aim is greater formalized achievement expectations, the results are apt to be disappointing. However, in line with the sound educational principle that learning should be satisfying, there is inestimable value in an educational climate which permits the expression of idiosyncratic needs, maintains a tolerance for innovation, is receptive to divergent-thinking activity, and encourages creative leadership. Such an atmosphere would be psychotherapeutic and growth facilitating, enabling significant development along such behavioral continuums as: rigidity—flexibility; unproductivity—fluency; unimaginativeness—fantasy; generality—originality. This approach is consistent with the idea of teaching the *whole* child, of meeting the global needs of the student, and of preserving the identity of the individual.

The cultivation of creative potentialities has heretofore been largely neglected by educators. If the ultimate goals are increased capability for personal expression, greater inventiveness, and the blossoming of gifted leaders,

the experiences may prove richly rewarding. It follows that out of the dedicated efforts to foster creative talent will come the development of achievement measures specifically designed to evaluate progress in learning to create and think productively. These *divergent achievement* indices would necessarily be based upon empirical developmental expectations regarding each of the significant dimensions of creativity. It is proposed that just as *IQ* is related to convergent achievement, an analogous relationship exists between *CQ* (Creativity Quotient) and divergent achievement.

E. SUMMARY

The comparative effects of exceptional intellectual potential and unusual creative aptitude on academic achievement were investigated. Four groups were constituted by levels of intelligence and creativity; respective subsamples represented children gifted in either of the two, in both, or in neither variable. Results of a two-way analysis of variance consistently demonstrated the significant role of intelligence in scholastic performance, while creativity was not determined to be related to academic success. The extent of general and test anxiety were also assessed, and found to be unrelated to intellectual ability or productive thinking. A factor analysis revealed the multifactorial nature of the creativity battery administered in this study; the loadings further described the various divergent-thinking subscales as bearing little relationship to each other.

Further exploration of the interaction between personality dynamics and creative activity is indicated. The dubious validity of measures of creative productivity suggests the need for establishing rigorous scientific criteria to preclude the spurious identification of "creative" individuals. Efforts should be made to distinguish between convergent and divergent achievement. Future research should focus upon those individuals with concomitant ability of a high order in both intelligence and creativity; it is this group which is perhaps truly representative of genuine giftedness.

REFERENCES

1. FLEMING, E. S., & WEINTRAUB, S. Attitudinal rigidity as a measure of creativity in gifted children. *J. Educ. Psychol.*, 1962, **53**, 81-85.
2. GETZELS, J. W., & JACKSON, P. W. The study of giftedness: A multidimensional approach. *The Gifted Student*, U.S. Department of Health, Education, and Welfare, Cooperative Research Monograph No. 2., 1960, 1-18.
3. ———. *Creativity and Intelligence: Explorations with Gifted Students*. New York: Wiley, 1962.
4. GUILFORD, J. P. Creativity. *Amer. Psychol.*, 1950, **5**, 444-454.
5. ———. The structure of intellect. *Psychol. Bull.*, 1956, **53**, 267-293.

6. ———. Three faces of intellect. *Amer. Psychol.*, 1959, **14**, 469-479.
7. HOUSTON, J. P., & MEDNICK, S. A. Creativity and the need for novelty. *J. Abn. & Soc. Psychol.*, 1963, **66**, 137-141.
8. KLAUSMEIER, H. J., HARRIS, C. W., & ETHNATHIOS, Z. Relationships between divergent thinking abilities and teacher ratings of high school students. *J. Educ. Psychol.*, 1962, **53**, 72-75.
9. KUBIE, L. S. *Neurotic Distortion of the Creative Process*. Lawrence, Kansas: Univ. Kansas Press, 1958.
10. LORGE, I. *Semantic Count of the 570 Commonest English Words*. New York: Teachers Coll., Columbia Univ., Bureau of Publications, 1949.
11. MACKINNON, D. W. The nature and nurture of creative talent. *Amer. Psychol.*, 1962, **17**, 484-495.
12. SARASON, S. B., DAVIDSON, K. S., LIDTHALL, F. F., WAITE, R. R., & RUEBUSH, B. K. *Anxiety in Elementary School Children*. New York: Wiley, 1960.
13. TAYLOR, C. W. The identification of creative scientific talent. *Amer. Psychol.*, 1959, **14**, 100-102.
14. THORNDIKE, R. L. The measurement of creativity. *Teachers Coll. Rec.*, 1963, **64**, 422-424.
15. TORRANCE, E. P. Educational achievement of the highly intelligent and the highly creative: Eight partial replications of the Getzels-Jackson study. *Res. Memorandum*, BER-60-18, Bureau of Educational Research, Univ. Minnesota, 1960.
16. ———. *Guiding Creative Talent*. Englewood Cliffs, New Jersey: Prentice-Hall, 1962.

Department of Psychological Services

East Williston Board of Education

East Williston, Long Island, New York

HAS THE STATUS OF MUSIC CHANGED IN 30 YEARS?*

Department of Psychology, Stanford University

PAUL R. FARNSWORTH

A. INTRODUCTION

It is a matter of common observation that words and sentences often change their meanings over time. This infiltration of new meanings has long plagued historians and linguists, scholars who can easily demonstrate that changes do occur but who can rarely if ever describe the extent of the changes. However, with the advent of scaling techniques, it has become possible to quantify semantic changes, at least in some small degree. The Thurstone technique of attitude scale construction (4), for example, provides for the assessing of the semantic values of scale items, i.e., allows for the sorting of attitudinal statements on a continuum, generally one with eleven steps. Thus, when college students were asked in 1930-31 where the statement "International disputes should be settled without war" should be placed on a continuum ranging from 1, expressing highest possible pacifism, to 11, expressing highest possible militarism, the median of their sortings came to 3.7. Further queries in subsequent years were made to demonstrate whether the item's value had remained unchanged or had moved toward the militaristic or the pacifistic end of the continuum (1, 2).

It should be noted that the sorters' judgments mirror their own attitudes only slightly; rather, the sorting medians reflect the prevailing attitudes found generally in the mores. Students from other colleges, and even reasonably literate noncollege adults, will sort items in an almost identical manner. In other words, other samples of sorters drawn at the same time from a roughly similar culture group will yield sorting medians quite similar to those given by the initial sorters.

Although at any one time the assessed semantic values of attitudinal items remain relatively constant within any culture group, they may change appreciably over the years. Thus, the median sorting value of the item mentioned above had moved from 3.7 to a more pacifistic 2.4 by the end of ten years. Suppose now that many other statements regarding war, sorted in 1930 and again after ten years, possessed median values which had also

* Received in the Editorial Office on May 31, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

moved toward the pacifistic end of the continuum. Then it could reasonably be concluded that the mores had shifted somewhat toward a more militaristic position. That is, if words and sentences come to seem more pacifistic than they formerly did, the point of reference must surely have become more militaristic.

B. METHOD

The present study follows an identical line of reasoning. The *Tests for Attitude Toward Music*, originally standardized in 1933-34 by R. H. Seashore and Kate Hevner (3), were restandardized in 1962-63. The median sorting values which emerged in 1962-63 were compared with those reported at the earlier date to learn the direction and magnitude of the semantic shifts. It is argued that a preponderance of large shifts in values toward the hostile-to-music end of the sorting continuum, if such occurred, should mean that the status of music is now better than it was 30 or so years ago, at least in the culture from which the sorters were drawn.

One hundred twenty-nine Stanford University students (45 were women) were taken from two large courses in elementary psychology and were asked to sort into 11 piles the 50 statements which made up the two forms of the Seashore-Hevner scale. A sorting into pile 1 meant that the judge regarded the statement as showing the most favorable attitude toward music possible, a sorting into pile 6 a neutral flavor, and a sorting into pile 11 the most hostile attitude. The semantic value of any one item was the median of the sortings. The Thurstone system assumes that pile 1's interval is from 0 to 0.99, 2's is from 1 to 1.99, etc. Hence, the medians may range from below 1 to above 10.

The medians were calculated in the traditional statistical manner except for the four items with values above 10. In figuring these, the Thurstone technique demands the use of extrapolations. However, since extrapolations with data of this sort tend to bring in rather subjective elements, it seemed best to ignore these four items and devote attention to the remaining 46. Yet it should be noted that the restandardization gave all four of these most hostile items values in excess of 10, showing little or no value changes between standardizations. Since the variability measures obtained in 1933-34 were no longer available, it was assumed that they probably would differ but slightly from those of the present study and, so, the current variability figures were used with both sets of data.

C. RESULTS

A simple count of the value changes between the two standardizations showed only 3 items completely unchanged in scale value. However, as the table below indicates, many other items showed relatively little change.

| Amount of Change in Values | | | | | | | | | | | | | | | | |
|----------------------------|---|----|----|----|----|----|----|----|----|----|-----|-----|-----|-----|-----|-----|
| Am't. | 0 | .1 | .2 | .3 | .4 | .5 | .6 | .7 | .8 | .9 | 1.0 | 1.1 | 1.2 | 1.3 | 1.4 | 1.5 |
| No. | 3 | 9 | 9 | 7 | 3 | 0 | 5 | 2 | 1 | 1 | 2 | 2 | 0 | 0 | 1 | 1 |

| Direction and Significance of Changes | | |
|---------------------------------------|--------------------|------------------|
| | Now more favorable | Now more hostile |
| No. of Changes | 21 | 22 |
| Significant at .01 level | 6 | 5 |
| .05 to .01 level | 2 | 2 |

Of the five items whose scale values now without question (.01 level of significance) are regarded as expressing more hostility to music than was true in 1933, one displayed the 1.5 value change, one the 1.4, two the 1.1 change, and one a 1.0 change. Of the six items whose scale values are now unquestionably regarded (.01 level of significance) as expressing more friendliness to music, the changes were 1.0, .9, .8, .7, .7, and .6. Four of the most significant changes in the hostile direction, then, were larger than any one of those which significantly changed in the friendly direction, and one was tied.

The item with the most dramatic shift (1.5) in scale value was "Music seems to have established itself as a part of our present day culture and I feel that it is an intellectual duty to keep in touch with it." In 1933-34 this item had a value of 3.3. While at the moment it is still regarded as a friendly item, its value has shifted to a more hostile 4.8. The item with the next largest value shift was "I enjoy songs where the words and music supplement each other, but pure music does not interest me." The earlier value here was 5.1 while now the value is 6.5. One of the items with a value shift of 1.1 (6.2 to 7.3) was "I like music but I could live the rest of my life without hearing any more of it." The other (4.5 to 5.6) was "If I thought that music could do for me what it seems to do for some people, I would be willing to spend more time on it." The fifth item whose value moved significantly (.01 level) in the hostile direction was "Most types of music are effeminate and have little possible interest for persons engaged in the more active pursuits of life" (8.8 to 9.8).

The six items whose values shifted most in the pro-music direction were: "I enjoy dance music and singing, but I am indifferent to the higher types of music" (6.0 to 5.0), "Although I do not spend a great deal of my leisure time on music, I would certainly not be willing to give up the time I do spend on it" (2.8 to 1.9), "To me there are few things which could be more bothersome than grand opera or symphony concerts" (9.6 to 8.8), "I should say that appropriate music always adds to my enjoyment of any drama or ceremony of any sort" (3.1 to 2.4), "The one or two real thrills that I get out of a concert are not worth the labor of listening to the whole concert" (7.8 to 7.1), and "From time to time I feel a need to listen to music or produce it myself" (2.8 to 2.2).

One can properly describe the above data by saying that, by and large, the statements in the Seashore-Hevner *Tests for Attitude Toward Music* have kept the affective flavors attached to them thirty years ago. Perhaps the best answer that can be given to the query posed in this paper's title is a qualified "No, the status of music has apparently not changed appreciably in 30 years." Yet the debater who might wish to uphold the position that music's status has shown at least slight improvement will receive a trifle more support from the data here reported than will his opponent. For the former can point out that the largest scale shifts favor his side of the argument.

REFERENCES

1. AYAD, J.-M., & FARNSWORTH, P. R. Shifts in the values of opinion items: Further data. *J. of Psychol.*, 1953, **36**, 295-298.
2. FARNSWORTH, P. R. Shifts in the values of opinion items. *J. of Psychol.*, 1943, **16**, 125-128.
3. MUELLER, J. H., MILL, E. G., ZANE, N. B., & HEVNER, K. Studies in appreciation of art. *Univ. Oregon Publ.*, 1934, **4** No. 6, 138-142.
4. THURSTONE, L. L., & CHAVE, E. J. *The Measurement of Attitude*. Chicago: Univ. Chicago Press, 1929.

Department of Psychology
Stanford University
Stanford, California

THE EFFECT OF PSYCHOPATHOLOGY IN CHILDREN'S DRAWING OF THE HUMAN FIGURE*

Department of Pediatrics, University Hospitals, State University of Iowa

THERON ALEXANDER¹

A. PURPOSE

Children's drawing of the human figure should be of especial interest to those studying the effects of psychopathology upon child development. Drawings may reveal distortions of perception or lack of attendance to essential parts of a Gestalt (8). It would be reasonable to expect that children with abnormal behavior would vary from normal children in the drawing of the human figure since their disorder might alter their perceptual experience. Children with morphological aberration in the central nervous system also might in turn differ characteristically from children with behavior disorder. In fact some investigators have suggested that definite characteristics of behavior are attributable to those children who have some deviation in brain structure. Does this deviation in structure have specific or characteristic effect on perceptual experience and does it affect performance on tasks requiring integration of perception into a Gestalt? For example, might such children tend to omit details of the face or might they concentrate on detail to the exclusion of essential parts of the figure? Now with increased knowledge of disorder and more precise diagnostic categorization, the study of children's abilities in various tasks may bring to light new evidence about characteristics of psychopathology. The specific questions to which answers are sought in this study follow. Do children with psychopathology fail entirely to achieve the Gestalt or totality of the human figure? Or, do differences exist in the use of details? And, further, are there patterns of omissions characteristic of pathology? Finally, what is the relationship of the use of parts to chronological age?

* Received in the Editorial Office on May 31, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ Acknowledgement is made of the collaboration in the diagnostic categorization of the subjects of R. B. Kugel, M.D., Director of the Child Development Clinic in the Department of Pediatrics. I wish to express appreciation of the interest and support in these studies of Donal Dunphy, M.D., Chairman of the Department of Pediatrics. I also appreciate the cooperation of Marjorie Clatterbaugh of the Iowa City public schools. Bruce Cushna, research associate, in the Child Development Clinic made valuable suggestions in the preparation of the manuscript.

This study seeks to answer these questions by examination of the use of perceptions of essential parts and of the integration of the parts into a whole. The answers are sought through (a) comparison of the performances in the use of essential parts of the human figure by normal children with the performances of those children who have some form of psychopathology, (b) the determination of the parts of the figure characteristically omitted by the normal and pathological groups, and (c) the exploration of the relationship of the number of parts of the figure to chronological age.

B. METHOD

1. Subjects

The 449 children in this study range in age from approximately six through twelve and are categorized as follows: (a) normal children, the entire population of an elementary school, (b) children with central nervous system (CNS) disorder (some form of encephalopathy), and (c) children with behavior disorder and no known physiological difficulties (with the exception of five who had psychophysiological disorders, but were kept in the behavior disorder group). The *IQ* means of the groups are: normal

TABLE 1
AGE MEANS AND RANGES FOR SUBJECTS

| Group | <i>N</i> | Mean | Range |
|-------------------------|----------|------|----------|
| Normal | | | |
| Male | 152 | 8.48 | 5.4-13.1 |
| Female | 123 | 8.84 | 5.5-13.0 |
| Total Normal | 275 | 8.64 | 5.5-13.1 |
| CNS disorder | | | |
| Male | 70 | 8.97 | 5.5-13.5 |
| Female | 36 | 8.75 | 5.5-13.8 |
| Total CNS disorder | 106 | 8.94 | 5.5-13.8 |
| Behavior disorder | | | |
| Male | 50 | 9.58 | 5.2-13.6 |
| Female | 18 | 9.33 | 6.6-12.4 |
| Total Behavior disorder | 68 | 9.51 | 5.2-13.6 |

(California Test of Mental Maturity) 124, behavior disorder (Stanford-Binet) 95, and CNS disorder (Stanford-Binet) 73. While one might argue that the groups should be matched on *IQ*, it was not done since the purpose of this study is to investigate the differences in the use of perceptions and the integration of these into a Gestalt. In other words, given the fact of disorder, what is the effect?

TABLE 2
SUBJECTS IN CNS DISORDER GROUP BY DIAGNOSTIC CATEGORIES

| Category | Number of Subjects | |
|---|--------------------|----|
| CNS disorder (encephalopathy) due to | | |
| Infection | 17 | |
| Prenatal | | 3 |
| Postnatal | | 14 |
| Intoxication | 1 | |
| Toxemia | | 1 |
| Trauma | 38 | |
| Prenatal injury | | 7 |
| Mechanical injury | | 2 |
| Anoxemia at birth | | 19 |
| Postnatal injury | | 10 |
| Disorder of metabolism, growth, nutrition | 3 | |
| Hypothyroidism | | 3 |
| Unknown prenatal influence | 22 | |
| Cerebral defect, congenital | | 18 |
| Mongolism | | 1 |
| Other | | 3 |
| Unknown cause, structure | 25 | |
| Associated with prematurity | | 1 |
| Structure with uncertain cause | | 24 |
| Total all subjects | 106 | |

TABLE 3
SUBJECTS IN BEHAVIOR DISORDER GROUP BY DIAGNOSTIC CATEGORIES

| Category | Number of Subjects | |
|---|--------------------|--|
| Behavior disorders | | |
| Aggressive behavior | 6 | |
| Over-inhibition | 32 | |
| Negativistic behavior | 15 | |
| Acultural behavior | 10 | |
| Psychophysiological disorders | | |
| Gastrointestinal reaction | 4 | |
| Genito-urinary reaction | 1 | |
| Total all subjects in behavior disorder group | 68 | |

2. Procedure

As part of the psychological study to categorize the children who had disorder, drawings of a man were obtained according to the instructions devised by Goodenough. Psychological study consisted of observations of the children at play, psychological testing [Stanford-Binet, Vineland Social Maturity Scale, and the Behavioral Complexity Test (1)] and parent interviews. Medical study included a medical history of pregnancy, birth, and development, medical examinations, and electroencephalograms.

For this study the human figure consists of clothing and 10 essential body

parts (head, neck, trunk, arms, legs, eyes, nose, mouth, hair, and ears.) The total number, 11, of these items provides a score for the study of the essential parts used. It is thus possible to determine not only the number of essential parts included, but their identity as well.

Goodenough's system for obtaining scores to determine *IQ* includes scorer judgments not only on the use of these parts and others, but also on proportion, coordination, and the like. The groups in this present study were also compared on her system of scoring.

3. Reliability

The reliability of the scoring methods used in the investigation was determined by correlation of the scores of 25 subjects in the normal group. The interscorer reliability obtained for two scores was .94.

C. RESULTS

For children in middle childhood, whether they have psychopathology or not, they, as grouped in this study, attain the Gestalt of a human figure in their drawings. The task, consequently, is well within their capacity. The groups, however, do differ significantly in the number of essential parts used. The normal children use the largest number of the essential parts as indicated by the mean of 10. Children with behavior disorder are next in order of complexity in their performance and attain a mean of 9 while children with

TABLE 4
ANALYSIS OF VARIANCE OF GOODENOUGH MODIFIED SCORES FOR NORMAL, BEHAVIOR
DISORDER, AND CNS DISORDER GROUPS

| Source | <i>df</i> | <i>ms</i> | <i>F</i> |
|--------|-----------|-----------|----------|
| Groups | 2 | 218.7 | 87.45* |
| Within | 446 | 2.5 | |
| Total | 448 | | |

* Significant at .01 level.

TABLE 5
MEAN DIFFERENCES ON GOODENOUGH MODIFIED SCORES

| Group | Mean | <i>t</i> |
|-------------------|------|----------|
| Normal | 9.91 | |
| Behavior disorder | 8.90 | 5.97* |
| Normal | 9.91 | |
| CNS disorder | 7.54 | 13.26* |
| Behavior disorder | 8.90 | |
| CNS disorder | 7.54 | 4.15* |

* Significant at .01 level.

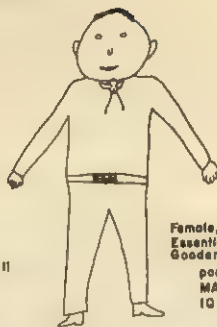
NORMAL SUBJECTS



Male, age 8yrs, 11mos
Essential parts score 8
Goodenough
point score 26
MA 9yrs, 3mos
IQ 104



Male, age 9yrs
Essential parts score 11
Goodenough:
point score 31
MA 10yrs, 6mos
IQ 117



Female, age 9yrs, 6mos
Essential parts score 10
Goodenough:
point score 31
MA 10yrs, 6mos
IQ 111

BEHAVIOR DISORDER SUBJECTS



Male, age 9yrs, 9mos
Essential parts score 9
Goodenough:
point score 28
MA 9yrs, 3mos
IQ 98



Male, age 9yrs, 10mos
Essential parts score 11
Goodenough:
point score 35
MA 11yrs, 6mos
IQ 117



Female, age 10yrs
Essential parts score 11
Goodenough:
point score 30
MA 10yrs, 3mos
IQ 103

CNS DISORDER SUBJECTS



Female, age 7yrs, 5mos
Essential parts score 7
Goodenough
point score 7
MA 4yrs, 9mos
IQ 84



Female, age 7yrs, 9mos
Essential parts score 7
Goodenough:
point score 10
MA 5yrs, 6mos
IQ 71



Male, age 9yrs
Essential parts score 9
Goodenough:
point score 18
MA 7yrs, 2mos
IQ 80

FIGURE 1
EXAMPLES OF DRAWINGS FROM GROUPS STUDIED

CNS disorder use fewest of the essential parts as shown by the approximate mean of 8. These means are high in terms of the 11 possible total points and, therefore, provide evidence that, generally, the children use most of the essential parts of the figure. These means also indicate that the subjects tend in all groups to use both gross parts of the figure, head, trunk, arms, and legs as well as the smaller details of the face.

Tests by the use of the chi square indicate that on the basis of presence or absence of essential parts used, the groups were not significantly different in the use of any one of the essential body parts. No one body part at all age levels consistently or characteristically was omitted by one group more than the other. One item included in the Gestalt of the human figure, clothing, however, did provide a significant difference among the groups. The use of clothing differentiated among the groups at all age levels.

At advancing age levels, the children in all groups increased their use of the essential parts and, accordingly, scores obtained on the essential parts correlated significantly with age. In order to determine if Goodenough's system also provides similar results, her total point system was correlated

TABLE 6
PEARSON CORRELATIONS BETWEEN AGE AND SCORING SYSTEMS

| Score type | Normal | Group Behavior disorder | CNS disorder |
|----------------------------|--------|-------------------------------|-----------------|
| Modified (essential parts) | .49** | .42** | .50** |
| Goodenough total points | .75** | .59** | .56** |
| Goodenough <i>IQ</i> | -.43* | -.32** | -.21* |

* Significant at .05 level.

** Significant at .01 level.

with chronological age. Her system, too, resulted in positive and significant correlations with age. Goodenough's system of obtaining *IQ*, however, did not provide similar results since significant negative correlations were obtained between the Goodenough *IQ* and chronological age.

D. DISCUSSION

Attainment of the Gestalt of the human figure is within the capability of most children in middle childhood in both types of disorder studied here. In the attainment of the Gestalt, both gross body parts and face details are used and there is no pattern of consistent omission of one part more than another. It seems, therefore, that a tendency to attend to one body part more than another is not generally found. Accordingly, influence of factors causing

group differences in the use of, or omission of, one body part does not occur. The use of clothing, however, does seem to be associated with diagnostic classification since this one factor is significantly different in the groups at each age level. Just why this differentiation should exist is unclear. Social class might be a factor, but the children were from varying social classes, and further, Britton (5) did not find a significant correlation between social class and Goodenough *IQ*. On the other hand, a study by Havighurst *et al.* (7) led these investigators to believe that culture produced differences in the drawing of the human figure. Whatever the basis is, clothing is the only consistent perceptual difference associated with psychopathology as it exists in these groups.

As children advance in age more parts of the human figure are used as has been well established in other studies. This finding is observed here and it is clear that as children grow older more use is made of perceptual experience, and expression in such tasks as drawing of the human figure reflects increased attention to essential parts contributing to the Gestalt. The findings in this study of the significant negative correlations of *IQ* with age suggest, however, that for some reason there is a problem in Goodenough's technique of obtaining mental age equivalents. Perhaps the experience of children has changed since her norms were established or perhaps the sample in this study is different in some way. The results obtained here do suggest that as children increase in age, the Goodenough norms for *IQ* represent some problem.

In summary, while one might expect brain pathology or behavior disorder to prevent integration of perceptions into a Gestalt, such is not the case. This study indicates that children with psychopathology differ from normal children in the number of parts used, but they can attain the Gestalt of the human figure and can depict both the gross parts and the face details. The characteristic use of only one essential part paralleled the diagnosis of pathology; the use of clothing differentiated among the groups at all age levels with the children having CNS disorder using clothing least and the normal children using it most.

The present study of children's drawings of the human figure is one of a series concerning a concept of response complexity. The concept, briefly described, is that interaction with the environment consists of perception of content of the environment, the accounting for its being, and the prediction of future events. In effective behavior, accordingly, there are the steps of perceiving condition, reasoning about causation (association with experience), and through further reasoning, prediction of future events or responses. The

broad purpose of these studies is to develop a concept of response complexity applicable both to normal and disordered individuals, to use the concept in understanding factors of effectiveness and pathology in behavior, to discover the relationship of the concept to development and age, to contribute to principles of learning or training, and to alleviate the deleterious effects of disorder.

This present investigation of drawing of the human figure relates only to the first phase, the perception of condition, and the integration of the perceptions into a unit of experience. The perceptions of the elements constituting a unit of experience are basic to effective response, and it is important to learn of the effects of pathology in this first phase. The task of drawing the human figure is not seen as reasoning about causation, nor as involving prediction about future response, but concerns only the first stage of the continuum of behavior described above. This investigation, therefore, concerns the use of the perceptions of essential parts, gross and detail, of the human figure. Earlier studies² of the first phase reported individual use of gross parts and details of the stimulus complex by subjects in response to test pictures. This present study, however, approaches the first phase of the behavioral sequence through the use of a different task, the drawing of the human figure.

Goodenough and Harris (6) published a comprehensive review of the studies done with the task of drawing the human figure and report that such studies date back for over a hundred years. Some of the previous studies are relevant to the present one. Kobayashi (10), Märtin (11), Katô (9), and Barnhart (2) indicate that with increasing age children tend to perceive the figure as a whole instead of a collection of parts. Bender (3) found that children who had had encephalitis, although performing adequately on other intelligence tests, still failed to do as well at the task of drawing the human figure. Berrien (4) found that those children who had had encephalitis (presumably with CNS disorder) failed to use the items of neck, hair, and clothing as frequently as did the children in the "psychopathic" and "mental deficiency" (borderline ability) diagnostic groups. Berrien also found that the children in the "psychopathic" group did not use clothing as an item as much as did the "mental deficiency" group. Although the diagnostic criteria

² The first study (Alexander, T., & Alpern, J. Psychological and physiological pathology of intellection. *Child Devel.*, 1960, **31**, 238) was reported at the 1959 meetings of the Society for Research in Child Development. The second study was reported at the 1961 meetings of the American Psychological Association (Alexander, T. The influence of central nervous system and behavior disorder upon complexity of response. *Amer. Psychologist*, 1961, **16**, 351).

are not clear, it seems that the postencephalitic and mental deficiency groups might be similar to the CNS disorder group used in the present study and that the psychopathic group was probably similar to the behavior disorder group presently used. If these suppositions are true then there is some contradiction in the findings since the present study indicates no consistent difference among the groups in body parts. And also in the present study the behavior disorder group used clothing more frequently than did the CNS disorder group. Comparison of findings is of limited value because of probable differences in diagnostic criteria.

The preceding studies of this investigation showed that subjects in groups with the same diagnostic classifications used in this present study did not differ generally on parts of pictures perceived, but differences lay in reasoning about the stimuli and in ability to predict future events or appropriate responses. There is, therefore, increasing evidence in view of present findings that effectiveness in behavior does not lie primarily in the perception of the existing conditions since in the drawing of the human figure most of the essential parts were perceived and no consistent omission of essential body parts was found; and, as was indicated in earlier studies, inadequacy lies further in the behavioral response pattern and reasoning about the stimuli perceived. The task of drawing of the human figure is, accordingly, considered to be a limited indicator of effective behavior and does not provide information about complex reasoning, but only perception and integration of parts into a Gestalt. Children studied attained the Gestalt and did not differ significantly in the exclusion of any one body part. It is concluded, thus, that important differences in intellection between those with psychopathology and those who are normal must be sought in more complex behavior.

REFERENCES

1. ALEXANDER, T., KUGEL, R. B., CUSHNA, B. & SNIDER, B. Studies of complex behavior: The processes of perception, association, and prediction of response. In press.
2. BARNHART, E. N. Developmental states in compositional construction in children's drawings. *J. Exper. Educ.*, 1942, **11**, 156-184.
3. BENDER, L. The Goodenough test (drawing a man) in chronic encephalitis in children. *J. Nerv. Ment. Dis.*, 1940, **91**, 277-286.
4. BERRIEN, F. K. A study of the drawings of abnormal children. *J. Educ. Psychol.*, 1935, **26**, 143-150.
5. BRITTON, J. H. Influence of social class upon performance on the Draw-a-Man Test. *J. Educ. Psychol.*, 1954, **45**, 44-51.
6. GOODENOUGH, F. L., & HARRIS, D. B. Studies in the psychology of children's drawings: II. 1928-1949. *Psychol. Bull.*, 1950, **47**, 369-433.

7. HAVINGHURST, R. J., GUNTHER, M. K., & PRATT, I. E. Environment and the Draw-a-Man Test: The performance of Indian children. *J. Abn. & Soc. Psychol.*, 1946, **41**, 50-63.
8. HENRY, W. E. *The Analysis of Fantasy*. New York: John Wiley & Sons, 1956.
9. KATÔ, M. A genetic study of children's drawings of man. *Jap. J. Exper. Psychol.*, 1936, **3**, 175-185 (*Psychol. Abst.*, 1937, **11**).
10. KOBAYASHI, S. A study on a variation of facsimilies drawn by children. *Jap. J. Psychol.*, 1937, **12**, 375-392 (*Psychol. Abst.*, 1938, **12**).
11. MÄRTIN, H. Die Motivwahl und ihr Wandel in der freien Zeichnung des Grundschulkindes (Choice of subject and its changes in the free drawings of elementary school children). *Z. Paedag. Psychol.*, 1939, **40**, 231-241 (*Psychol. Abst.*, 1940, **14**).

Department of Pediatrics
University Hospitals
State University of Iowa
Iowa City, Iowa

THE RETENTION OF SENSORY EXPERIENCE: I. INTENSITY*

University of Pittsburgh School of Medicine and Western Psychiatric Institute and Clinic

H. E. KING

A. INTRODUCTION

It has long been known that a human observer, acting under instruction, can compare a present sensory stimulus with one that has gone before in time, i.e., he is able to make a form of successive comparison of noncoincident sensory stimuli. In some way a "trace" left by a first stimulus can be mentally placed alongside a second, later stimulus and a judgment formed about their relative value (5, p. 440). Data generated by this technique have been used experimentally, particularly to probe into the time-errors that often appear in relating sequential percepts (3, 5), and have also served quite practical ends, as in the evaluation of the sense of pitch on tests of musical talent (4). Most of the uses made of the method of successive comparison have separated the standard stimulus (the first in time) and the variable stimulus (the second) by the briefest of delay intervals, by seconds or fractional seconds, and a few have extended the interstimulus interval to minutes or hours. Little is known about the consequences of interstimulus delays of greater length. The purpose of the present study, and those related to it to follow (1, 2), is to focus on the variable of prolonged delay between a standard and a variable stimulus presented in a context of successive comparison to observe, systematically, human ability to retain specific sensory impressions over periods of days, weeks and months. The experimental situation thus framed presents a cross between the psychophysical observation and a "memory" study in a more traditional sense. It should provide information on human ability to retain relatively raw sensory experiences, or simplified percepts, that are devoid of obvious meaning, association or symbolic reference and may be expected to supplement what has been learned about human storing capacities using non-sense verbal materials.

B. METHOD

Vision and hearing have been selected as the most convenient modalities with which to begin exploration and for the degree of precision in measure-

* Received in the Editorial Office on May 31, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

ment that they permit. The test procedures are a form of the method of adjustment, where a second stimulus (the *variable*) is adjusted by manual control to a value judged to be equal to the stimulus that has gone before (the *standard*). Each subject was exposed to a stimulus of known physical character and, after a predetermined delay, was asked to approximate his "memory" of this earlier stimulus value. All changes made in the variable stimulus were effected by rotary manual controls that provided a wide range of possible response within a single dimension, here intensity.

1. *Visual Intensity (Brightness)*

All observations were made with the subject seated in a darkroom with 1.0 mL. constant ambient illumination, measured from a true white surface normal to the line of regard to the test apparatus. Ten minutes of adaptation to this illumination level elapsed during general instruction. The subject was told that, following a ready signal, a window on the test panel before him would be illuminated for five seconds by a white light, and that he was to inspect the field and notice how bright it was; following a delay, he would be asked to reproduce the degree of brightness of the original stimulus—as best he could approximate it—by rotation of a geared control knob that would vary the intensity of the stimulus field.

A standard stimulus brightness of 74.5 mL. was presented to all subjects. The test field was circular, 32 mm. in diameter, placed 22 inches from the corneal surface and viewed horizontally with both eyes; a visual angle of 3.3°. Light from an incandescent source, powered by a monitored transformer, passed through a stationary and a movable Polaroid filter and opal and translucent glass diffusing layers. The position of the movable Polaroid filter, circumference mounted to a geared rotation drive, could be read to an accuracy of one-half degree of arc of rotation. A range of field brightness from 5–170 mL. was possible by adjustment of relative Polaroid position; all brightness determinations were made by Macbeth illuminometer.

2. *Auditory Intensity (Loudness)*

All observations were made with the subject seated in a sound-deadened chamber, with an ambient noise level of 20–35 db., measured at the point of the subject's right ear. The subject was told that, following a ready signal, he would hear a tone for five seconds and he was asked to note how loud it was; that upon return he would be asked to duplicate this loudness by adjustment of a manual control that would vary the intensity of the same stimulus tone.

A standard stimulus tone of 600 c.p.s., generated by a General Radio Beat-Frequency Audio Generator, was sounded for all subjects at an intensity of 43.8 db., measured at the subject's right outer ear. Both ears were used to listen to sound stimuli delivered from an AR-2 acoustic-suspension speaker; ear to speaker distance was 56 inches in a lateral horizontal plane. A range of sound intensity from 0-100 db. was possible by rotation of a micropotentiometer in the output of the oscillator; all intensity determinations were made by a General Radio Sound Level Meter (re 0.0002 microbar).

C. SUBJECTS AND PROCEDURE

Five groups of subjects, student nurses, were observed; 16 in each group. One-half of each subject group made a match to a "trace" but once, after a prolonged delay; the other half made one such match at the same sitting, after a two-minute delay, were exposed to the standard again for five seconds and made a second match after a prolonged delay. Random sequences were followed in determining the factors of: test order, delay interval, stimulus value from which reset begun and whether the subject made both a two-minute reset and a delayed match, or a delayed match only. The intervals observed between exposure to the standard and resetting of the variable stimulus were two minutes and 1, 7, 14, 21 and 28 days. A constant monitoring of the test apparatus, by electric and photometric standards, was required to maintain the physical values for the stimuli at consistent levels over the extended experimental period.

D. RESULTS

1. *Visual Intensity (Brightness)*

The intensity of the standard stimulus and the mean intensity for all delayed resettings of the variable stimulus are given in Table 1, and have been graphed for visual inspection in Figure 1. All of the matches made to the standard brightness show a close approximation to its known value. The full range of response possible (5-170 mL.) is not shown on the ordinate of Figure 1, and it may be noted that mean reset values fall within 20 mL. of the correct (standard) value at all intervals of delay. There is a slight reduction in the mean reset value, not significant by *t* test, for matches made at the same sitting, two minutes after exposure to the standard. Over the period of longer delays, ranging from one to 28 days, a similar slight reduction below the standard is observable, but a middle estimate of performance by these independent groups (M) does not differ reliably either from the value of the standard or of resettings made following only a two minute delay.

TABLE 1
INTENSITY VALUES FOR THE STANDARD STIMULUS AND
MEAN RESET PERFORMANCE AFTER DELAY

| | Standard stimulus | 2 min. | 1 day | Variable stimulus | | 21 days | 28 days |
|---------------------|-------------------|--------|-------|-------------------|---------|---------|---------|
| Brightness (in mL.) | 74.5 | 64.8 | 70.8 | 7 days | 14 days | 58.8 | 75.5 |
| Loudness (in db.) | 43.8 | 47.5 | 50.3 | 51.2 | 55.7 | 54.7 | 53.6 |

An analysis of variance was made to intercompare matching performance by the subject groups for the longer delay periods (1-28 days) and to contrast the conditions of resets begun at intensities brighter, and less bright, than the standard, and for subjects making a single match with those making one

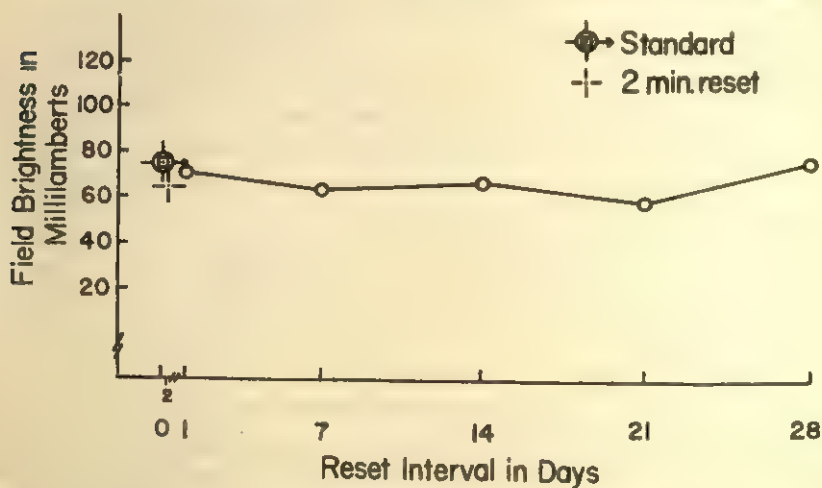


FIGURE 1
BRIGHTNESS OF THE STANDARD VISUAL STIMULUS AND MEAN RESETTINGS
OF THE VARIABLE STIMULUS AFTER DELAY

reset at the same sitting (two-minute delay) and again after a prolonged delay (1-28 days). All F ratios computed were nonsignificant for main conditions. One interaction, *delay* with *direction*, was significant; $P < .05$. The group variation remained relatively constant at all delay intervals.

2. Auditory Intensity (Loudness)

The intensity of the standard stimulus and the mean intensity for all delayed resettings of the variable stimulus are given in Table 1 and have

been graphed for visual inspection in Figure 2. The matches made to the standard loudness show, on the whole, a reasonable approximation to its known value at each delay interval. The full range of response possible (0-100 db.) is not shown on the ordinate of Figure 2 and it may be noted that mean reset values fall within 12 db. of the correct (standard) value at all intervals of delay.

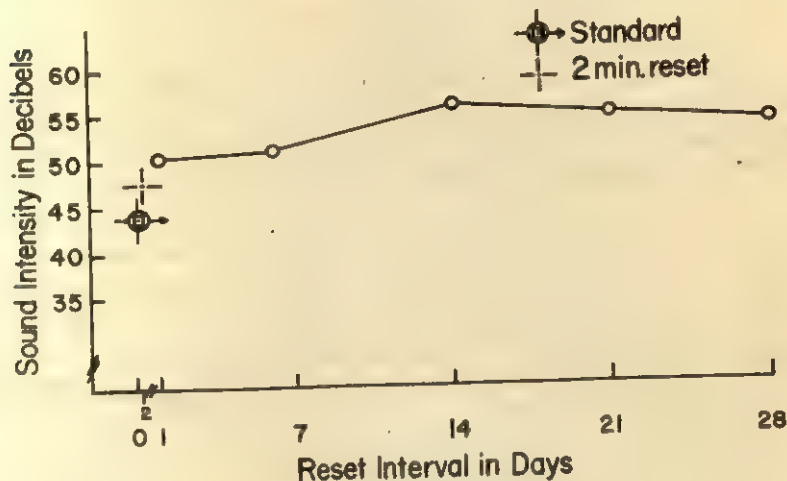


FIGURE 2
LOUDNESS OF THE STANDARD AUDITORY STIMULUS AND MEAN RESETTINGS
OF THE VARIABLE STIMULUS AFTER DELAY

There is an observable increase in the mean intensity matches made at the same sitting, after a two-minute delay, that is significant by t test; $P < .01$. Over the period of longer delays, from 1-28 days, further increases in intensity resettings can be noted, a middle value of which differs reliably from the standard; $P < .01$, and also from the mean of resettings made following a two-minute delay; $P < .01$.

An analysis of variance was made to intercompare matching performance by the subject groups for delay periods of 1-28 days, and to contrast the conditions of resets begun at intensities louder, or less loud, than the standard and for subjects making only a single match with those making one reset at the same sitting (two-minute delay) and again after a prolonged delay (1-28 days). Significant F ratios were obtained for *delay*, $P < .05$ and for an interaction of *delay* with *direction*, $P < .05$. The group variation remained relatively constant over all delay intervals.

E. DISCUSSION

It would appear, from these findings, that naive subjects are able to reproduce quite faithfully the intensity of a standard visual or auditory stimulus to which they have been exposed but briefly, whether the delay between viewing the original and matching its "trace" is as short as two minutes or as long as 28 days. The percepts formed of the strength of the original stimuli seem to be sufficiently long-lived to permit a much delayed, successive comparison of remarkable accuracy. Just what it is that is carried through time, here called a "trace," is not made more explicit by these data, but the ability of human observers to bear in mind what Woodworth has called "the effective standard," to which new stimuli may be related after prolonged delay, is demonstrated and displays a decided degree of regularity.

Beyond the indications of the stability of the "traces" so borne over minutes, days or weeks is the apparent contrast in the performance of subjects on visual and auditory sample tasks. Although the successive matches for both depart from the value of the standard on resettings made after a two-minute delay, the difference was a significant one for the auditory matches, but not for the visual. Pursuing the comparison by modality through performance at the longer delay intervals, it appears that while matches made to the intensity of the visual stimulus never depart significantly, either from the standard value or from the resettings made within two minutes, the intensity matches to the auditory stimulus deviate further, and significantly, from the standard value, from the resettings made after a two-minute delay and still further as the interval extends through the 28-day period of observation. The "trace" for visual intensity, at least for this sample task, appears to be more lasting and more stable than the aftereffect of auditory intensity.

Neither the comparison of performance by subjects beginning their search for a matched intensity value from a level more intense, or less intense, than the standard, nor a comparison of performance by those who made a single match with those making two, produced a significant difference in the resettings observed for sample tasks on either sense modality. A non-significant trend in the data that can be seen in the former condition, however, deserves brief comment. It is known that many judgments in psychophysics display a "direction" effect, i.e., there can be an influence exerted on a judged point by whether the approach to that point originates from stimuli "above" or "below" it, as is commonly reported for data collected by the Method of Limits. A similar effect appears in the adjustments made in

these experiments, by successive comparison, where the resetting judged to be a match to the standard tends to be displaced in the direction from which the search for the judged point was begun (more intense or less intense). Although the direction of search was balanced in the design of the experiments and the difference in resettings made fell short of statistical significance when analyzed as a condition, it does appear to combine with the effect of delay to produce significant interactions for both visual and auditory intensity matchings. The same trend may provide, in part, an explanation for the relatively constant group variation obtained at all delay intervals by tending to expand the group variability. Consistency in group variation might also be expected in view of the relative constancy of mean performance over the differing delay intervals. Additional experiments that are now in progress center on this factor and may permit a more accurate weighing of the overall influence of a "direction" trend on resetting performance.

All of the subjects were able to grasp what was expected of them easily and showed a general confidence in being able to make such a match-through-time, especially for the shorter delay intervals. Volunteered introspective comment indicated that subject confidence diminished for matches made after the longer periods of delay, despite a generally adequate performance. Each subject may view the experimental task as a form of figure-on-ground impression and will, of course, relate the sample intensities of the experiment to his entire realm of sensory experience. This process itself invokes a judgment of successive comparison among "traces." It should be stressed that half of each group had no experience whatever with the range of stimulation possible in the experimental situation nor any knowledge of the physical character of the adjusting controls; the remaining half, with one opportunity to explore the range fully and systematically, rarely did so and did not differ in the intensity matches chosen to either the sample visual or auditory standard stimulus. To compare the subjects' matching performance on differing sense modalities, across qualities, at least rough comparability of standard stimulus levels should obtain. An attempt has been made to meet this requirement by selecting as standards intensities falling slightly below the half-range point for each sense, where neither a true zero nor extreme intensities were obtainable by the subjects' adjustment of the variable stimulus.

The subjective approximations made to "traces" of the intensity characteristic of selected auditory and visual stimuli indicate, on the whole, an impressive stability among different subject groups over delay periods of minutes, days and weeks; show an interesting relation between the deviations

made at the same sitting and those made after much greater delay and suggest the presence of genuine differences in subject response to stimuli in the visual and auditory sense modalities. Additional experimental observations made upon the same sense modalities, but exploring different stimulus dimensions, will permit a more rounded inspection of human ability to retain simple sensory percepts over prolonged intervals of time (1, 2).

F. SUMMARY

The accuracy of human subjects in retaining an experience of sensory intensity over long delay intervals was observed. Standard stimuli, presented briefly in the visual and auditory modalities, were matched by a manual adjustment of intensity controls following a delay of minutes, days or weeks. The approximations made to "traces" of the intensity dimension of selected stimuli demonstrated a remarkable stability over long delay periods, evanescent only slowly, or not at all, over a one-month period of observation. Greater changes were observed for matches made to auditory than to visual stimuli; matches for both senses showed the largest shifts in judged intensity equivalence after the shortest delay.

REFERENCES

1. KING, H. E. The retention of sensory experience: II. Frequency. *J. of Psychol.*, 1963, **56**, 291-298.
2. ———. The retention of sensory experience: III. Duration. *J. of Psychol.*, 1963, **56**, 299-306.
3. OSGOOD, C. Method and Theory in Experimental Psychology. New York: Oxford Univ. Press, 1953.
4. SEASHORE, C., LEWIS, D., & SALTVEIT, J. Manual of Instructions and Interpretations for the Seashore Measures of Musical Talents (1938 revision). New York: Psychological Corp., 1938.
5. WOODWORTH, R. Experimental Psychology. New York: Holt, 1938.

Western Psychiatric Institute & Clinic
3811 O'Hara Street
Pittsburgh 13, Pennsylvania

THE RETENTION OF SENSORY EXPERIENCE: II. FREQUENCY*

*University of Pittsburgh School of Medicine and Western
Psychiatric Institute and Clinic*

H. E. KING

A. INTRODUCTION

The ability of human subjects to compare successive, as well as simultaneous, sensory stimuli has found use in the psychophysics laboratory and in practical contexts for many years, and is an observation commonplace enough to be a part of everyday experience; for example, in recognizing faulty machine-operation by a change in sound from its usual "hum" or in matching colors that are before the eye to a sample carried only in the mind. It has been proposed, in a recent paper, that controlled and detailed observation of the adequacy with which simplified sensory percepts can be carried through long delay intervals might provide useful information about basic "storage" capacities of the human organism and serve as a convenient and theoretically interesting source of nonverbal material with which to study the impact of time passage on specified experience (1). The first-reported experiments turned attention to the intensity characteristic of percepts formed in the visual and auditory modalities, observing the accuracy of human observers in finding again the strength of sample stimuli experienced but once—following delays varying in length from minutes to a month. This paper will explore the retention of the characteristic of stimulus frequency for sample stimuli in the same two modalities, following an experimental approach otherwise identical in outline.

B. METHOD

A form of the method of adjustment was used in collecting data. The subject was first exposed to a *standard* stimulus of known physical character, a delay was imposed and he then adjusted a *variable* stimulus, manually, to judged equivalence with his "memory" of the standard.

* Received in the Editorial Office on May 31, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

1. *Visual Frequency (Flash Rate)*

All observations were made with the subject seated in a darkroom with 1.0 m.L. ambient illumination, measured from a true white surface normal to the line of regard to the test apparatus. After 10 minutes of adaptation the subject was asked to view the panel before him where, following a ready signal, a window would be illuminated for five seconds by a flashing light. He was to inspect the field, form an impression of the rate of flash and, after a delay, be asked to reproduce this rate—as best he could approximate it—by rotating a geared control knob that would vary the frequency of flash of the test field.

A standard stimulus flash-rate of 25.0 c.p.s. was presented to all subjects. The test field was circular, 25 mm. in diameter, placed at a distance of 18 inches from the corneal surface and viewed horizontally with both eyes; a visual angle of 3.1° . Light from a gas discharge tube, powered by a General Radio Strobotac, Type 631-B, passed from the source through an opal glass diffusing layer, emerging from the test field at a brightness of 64 m.L. for the standard stimulus. The range of flash-rates possible was limited at the low end by the limit of the instrument and at the high end by the subject's critical fusion frequency for this stimulus area and intensity; approximately 10–50 c.p.s., with an accuracy of dial reading $< .2$ c.p.s.

2. *Auditory Frequency (Pitch)*

All observations were made with the subject seated in a sound-deadened chamber, with an ambient noise level of 20–35 db., measured at the point of the subject's right ear. Each subject was told that he would hear, after a ready signal, a tone sounded for five seconds and was asked to attend to its pitch; i.e., how high or low a note it was. On his return to the test chamber he would be asked to find this note again by the manual rotation of a control knob that would vary the pitch of the tone being sounded.

A standard stimulus tone of 1000 c.p.s., generated by a General Radio Beat-Frequency Audio Generator, was presented to all subjects at an intensity of 75.0 db., measured at the subjects' right outer ear. Both ears were used to listen to sound stimuli delivered from an AR-2 speaker; ear to speaker distance was 56 inches in a lateral horizontal plane. A range of audible frequencies from 20–20,000 c.p.s. was possible by rotation of the adjusting control, limited by the hearing characteristics of the individual listener. The energy output of the oscillator is near constant across its range, and an inference about heard loudnesses, based on equal-loudness contour curves,

indicates that tones near the test frequency vary little in intensity [< 4 db. over a range of 500–4,000 c.p.s. (2, p. 504)]. All determinations of sound intensity were made by a General Radio Sound-Level Meter (re 0.0002 microbar).

C. SUBJECTS AND PROCEDURE

Five groups of subjects, student nurses, were observed; 16 in each group. One-half of each group matched to a standard but once, after a delay of one or more days; one-half made one match to a "trace" of the standard after a two-minute delay, were re-exposed to the standard and made a second match following a delay of one or more days. Random sequences were followed in determining the factors of: test order, delay interval, stimulus value from which reset was begun and whether the subject made both a two-minute reset and a match after prolonged delay, or a single match only after prolonged delay. The delay intervals between exposure to the standard and attempts to reset the variable stimulus were: two minutes and 1, 7, 14, 21 and 28 days (1).

D. RESULTS

1. Visual Frequency (Flash-Rate)

The frequency of the standard stimulus and the mean frequency for all resettings of the variable stimulus after delay are shown in Table 1 and have been graphed in Figure 1. On the whole, variable matches made to the standard tend to approximate its known value, but with a consistent displacement toward rates of higher frequency. The full range of response possible (10–50 c.p.s.) does not appear on the ordinate of Figure 1 and the

TABLE 1
FREQUENCY VALUES FOR THE STANDARD STIMULUS AND
MEAN RESET PERFORMANCE AFTER DELAY (IN C.P.S.)

| | Standard stimulus | Variable stimulus | | | | | |
|------------|-------------------|-------------------|-------|--------|---------|---------|---------|
| | | 2 min. | 1 day | 7 days | 14 days | 21 days | 28 days |
| Flash-Rate | 25.0 | 29.8 | 31.1 | 29.7 | 29.8 | 28.5 | 35.3 |
| Pitch | 1000 | 1774 | 1592 | 1863 | 1935 | 1521 | 1785 |

mean resettings of the variable stimulus fall within 10 c.p.s. of the standard at all intervals of delay. There is an obvious increase in mean reset value, significant by t test ($P < .01$), for frequency matches made at the same sitting, two minutes after exposure to the standard. Over the period of much longer delays, 1–28 days, a similar increase in judged rate may be observed for adjustments made of the variable stimulus, a middle value of which (M)

is also significantly higher than the standard flash-rate ($P < .01$), but does not differ reliably from resettings of the variable made after two minutes of delay.

An analysis of variance was made to intercompare matching performance by the independent groups over the period of delay from 1-28 days and to contrast the conditions of resets begun at flash-rates higher than, and lower

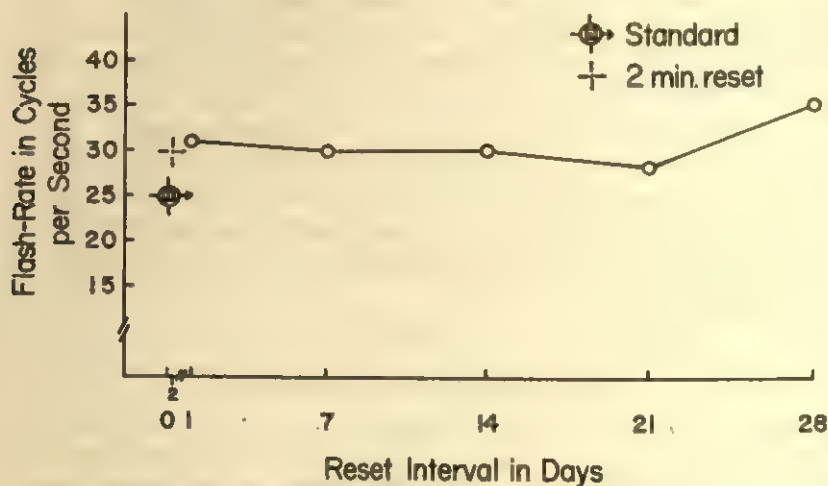


FIGURE 1
FLASH-RATE OF THE STANDARD AND MEAN RESETTINGS OF THE VARIABLE
STIMULUS AFTER DELAY

than, the standard rate and for subjects making a single match with those making one reset after a two-minute delay and again after a delay of one or more days. A significant F ratio was obtained for the main condition of *direction*, $P < .01$; none of the comparisons of other main conditions or interactions reached significance. Group variation remained approximately constant at all delay intervals.

2. Auditory Frequency (Pitch)

The frequency of the standard stimulus tone and the mean frequency for all delayed resettings of the variable stimulus are shown in Table 1 and have been graphed in Figure 2. Overall, the variable matches approximate the value of the standard stimulus, with a consistent tendency toward matching with tones of higher pitch. The full range of response possible (20-20,000 c.p.s.) does not appear on the ordinate of Figure 2 and all

mean values for adjustment of the variable stimulus fall within 1000 cycles of the standard at all intervals of delay. There is a visible increase in the mean reset value, significant by t test ($P < .01$), for matches made at the same sitting, two minutes after exposure to the standard. Over the period of longer delays, 1-28 days, a similar increase in judged equivalence of frequency may be observed for adjustments made of the variable stimulus, a middle estimate of which (M) is also significantly above the standard frequency ($P < .01$), but does not differ reliably from resettings of the variable stimulus made after two minutes of delay.

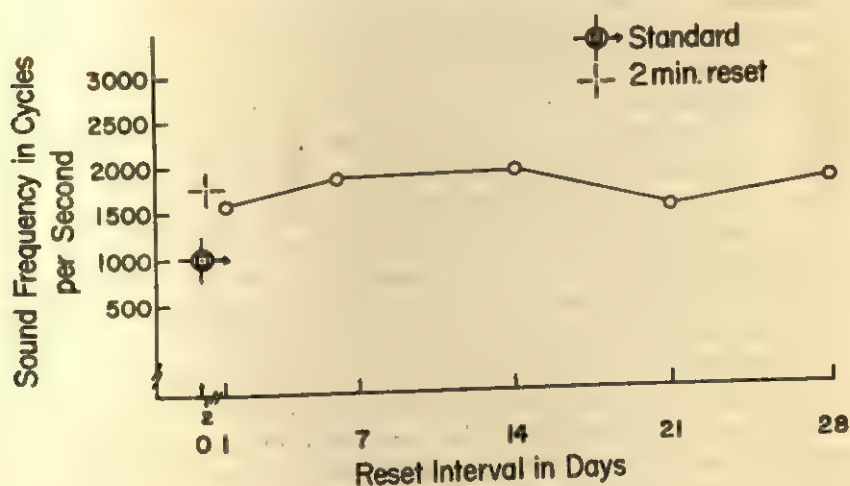


FIGURE 2
PITCH OF THE STANDARD AND MEAN RESETTINGS OF THE VARIABLE
STIMULUS AFTER DELAY

An analysis of variance intercompared performance, by the subject groups over the period of delays from 1-28 days, by whether the beginning variable stimulus was of higher or lower frequency than the standard and contrasted performance by those subjects making one match, after prolonged delay, with those making one match after a two-minute delay and, after a re-exposure to the standard, a second match after a delay of one or more days. Significant F ratios were obtained for comparisons of the main conditions of delay ($P < .01$) and direction ($P < .01$), and for the interaction of delay and direction, ($P < .01$). Group variation remained approximately constant at all delay intervals.

E. DISCUSSION

The adjustments made of a variable stimulus to match the "memory" of a prior sensory experience indicate that the subjects were generally able to approximate the frequency value of standard stimuli in both visual and auditory modalities. A clear tendency was expressed toward selecting matching stimuli of somewhat higher frequency than the standard for both senses, an effect that is manifest in judgments made after the shortest delay (two minutes) and which appears again in the resettings made after each of the longer delay intervals (1-28 days) without further increase. A degree of departure from the actual frequency of both visual and auditory standard stimuli was, then, immediately evident, but those successive-comparisons made after much longer periods of delay—a day, a week or a month—were no less accurate than were judgments made under more optimal seeming conditions, i.e., with short delay and at the same experimental sitting.

The form of the curves describing retention of the frequency characteristic of standard visual and auditory stimuli is quite similar, overall, and suggests no such contrast in performance between these same modalities as had been earlier noted for impressions carried through time of stimulus intensity (1). Although a significant difference did appear in performance over the 1-28 day delay period for judgments of auditory frequency and did not for visual frequency, an inspection of Table 1 and Figure 2 suggests this to be a difference between the highest and lowest mean frequency values chosen by these independent groups rather than denoting a progressive change or a systematic trend. The successive-comparisons to a frequency standard were also quite alike, for the two modalities, in the strong tendency displayed toward arresting judgment short of the actual standard frequency, whether the variable stimulus with which search for a match was begun was at a rate faster, or slower, than the standard stimulus. As the experimental design was balanced for the direction of search, these early-arrests in either direction tend to neutralize one another, as is done in combining the data of ascending and descending sensory thresholds by the Method of Limits. It is likely that the presence of a trend toward early-arrest in judged equivalence would expand the estimates made of group variance, however, and this may be a contributing factor to the relative stability of group variation observed in matching performance after all delay intervals. The generally stable mean performance over time by different subject groups also suggests that a consistency in group variation might be expected to occur.

No difficulty was expressed by the subjects in grasping what they were

expected to do in the test situation and a general air of confidence was shown in their potential ability to make such equatings of stimulus frequency through time. A greater uncertainty in actually making these judgments was manifested (e.g., by the time needed to complete, volunteered verbalizations, etc.) than had been shown for intensity matchings, however, both on those resetttings made at the same sitting and those attempted after prolonged delay. No difference was evident in the performance of those subjects making only a single match after long delay and those making one match at the same sitting and another after delays of one or more days for stimuli in either modality.

The ability of human subjects to find again the frequency rates of visual and auditory stimuli, to which attention had been directed during brief exposure to the standards, indicates, on the whole, that good approximations can be found whether the delay between original and variable stimulus is as short as two minutes or as long as 28 days. A constant error in selecting frequencies higher than the standard was noticeable after the shortest delay and again after each of the prolonged delays, but does not seem to be progressive. That is to say, the degree of error shown in successive-comparison to the frequency characteristic of sample visual and auditory stimuli was no greater after the lapse of weeks than it was after minutes. The presence of strong "direction" effects in the data suggests that method has its influence and interaction with what is observed and urges caution in generalizing from the results thus far obtained.

It would appear, from these findings, that naive human subjects are quite able to approximate the frequency characteristic of sample visual and auditory sensory experiences, and that little or no evanescence of "trace" was demonstrated with delays as great as four weeks elapsing between exposure to the standard and a judged equivalence of match by a variable stimulus. Taken in conjunction with the findings for the intensity characteristic of stimuli in the same modalities (1), it would seem that relatively simplified sensory percepts may be remarkably longlived in the human, where attention has been drawn in advance to the need to "bear them in mind."

F. SUMMARY

The accuracy with which naive human observers were able to reproduce an experience of sensory frequency after short and long intervals of delay was determined. Standard visual and auditory stimuli were presented at known frequency rates for a five-second period, then matched by a manual adjustment of a variable frequency control following delays of minutes, days

and weeks. The approximations made to "memory" of the frequency characteristic of the sample stimuli demonstrated remarkable stability, for both senses, over all delay periods. A constant error in the direction of raised frequency was noted in both modalities that appeared to be much the same after all intervals of delay. Judged equivalence of frequency in both modalities showed the greatest shift after the shortest delay interval.

REFERENCES

1. KING, H. E. The retention of sensory experience: I. Intensity. *J. of Psychol.*, 1963, **56**, 283-290.
2. WOODWORTH, R. S. *Experimental Psychology*. New York: Holt, 1938.

Western Psychiatric Institute & Clinic
3811 O'Hara Street
Pittsburgh 13, Pennsylvania

THE RETENTION OF SENSORY EXPERIENCE: III. DURATION*

*University of Pittsburgh School of Medicine and Western
Psychiatric Institute and Clinic*

H. E. KING

A. INTRODUCTION

The influence of time-passage on memory for sensory experience may be approached, experimentally, by making use of the human subject's evident ability to compare successive percepts and his skill in relating the value of a second sensory impression to another, like it, that has gone before. By varying the length of the interval between the original and subsequent experiences, and by turning attention to only one stimulus dimension at a time, systematic observations may be made of the accuracy with which impressions of the kind may be carried, or "stored," over periods of minutes, days, weeks or months. Two recent experiments have described the responses of human subjects to the stimulus properties of intensity and frequency for percepts formed by way of the visual and auditory modalities (4, 5). The focus of the present paper will be on the characteristic of stimulus duration, for sample stimuli briefly experienced in the same modalities, following an experimental approach otherwise parallel in general outline.

B. METHOD

The method of reproduction has been used in collecting data on the retention of brief stimulus durations. Each subject was first exposed to a *standard* stimulus of known duration, a delay was imposed and he then attempted to reproduce the original interval by pressing a reaction key to judged temporal equivalence with his "memory" of the standard. This form of time estimation gives the closest approximation available to the method of adjustment of the *variable* stimulus by manual control that was used in prior study of the intensity and frequency properties of simple percepts (4, 5). As time moves only forward, an adjustment from longer to shorter is not possible and the reproduction task thus becomes a one-ended adjustment, from zero to judged equivalence, of a *variable* temporal stimulus.

* Received in the Editorial Office on May 31, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

1. *Visual Duration (Flash Length)*

All observations were made in a darkroom with dim, ambient illumination of 1.0 m.L. measured from a true white surface normal to the line of sight to the test panel (4). The chamber was sound-deadened; the ambient noise level varying between 20–35 db. when measured at the point of the subject's right ear, with no entry of sound from periodic sources (5). After becoming accustomed to the chamber for 10 minutes, the subject was asked to attend to a window on the test panel that would, following a ready signal, be illuminated for a brief interval by a diffused violet light. He was told to inspect the field, form an impression of how long the light remained on and, after a delay, be asked to reproduce this flash-length—as best he could approximate it—by pressing a reaction key that would again illuminate the test field for an interval of his choosing.

A standard stimulus duration of 2.0 seconds, at a brightness of 2.7 m.L., was presented to all subjects. The test field was circular, 19 mm. in diameter, placed 22 inches from the corneal surface and viewed horizontally with both eyes; a visual angle of 1.9° . Light from an argon tube of rapid on-off character ($< .004$ seconds) passed through an opal glass diffusing layer forming the surface of the test field. The length of flash of the standard stimulus was controlled by electronic timer and the variable flash by the subject's closure of a telegraph key; 1 mm. contact gap at 200 gms. minimal pressure to close. The subject's response could vary from zero to infinity and was read, in .01 seconds, by electric chronometer.

2. *Auditory Duration (Tone Length)*

All observations were made in the same chamber and situation that is described above. The subject was told that, following a ready signal, he would hear a tone sounded and he was asked to note how long it remained on; that upon return he would be asked to duplicate this interval by pressing a response key that would allow the same tone to be heard again for an interval of his choosing.

A standard stimulus duration of 1.5 seconds was presented to all subjects. The tone, at 200 c.p.s., was generated by a General Radio Beat-Frequency Audio Generator at an intensity of 80 db. (re 0.0002 microbar), measured at the point of the subject's right outer ear. Both ears were used to listen to sound stimuli delivered from an acoustic-suspension speaker; ear to speaker distance was 56 inches, in a lateral horizontal plane. The standard tone length was controlled by an electronic timer and the variable tone by subject

closure of a telegraph key; 1 mm.-contact gap at 200 gms. minimal pressure to close. The subject's response could vary from zero to infinity, and was measured by electric chronometer in .01 seconds.

C. SUBJECTS AND PROCEDURE

Five groups of subjects, student nurses, were observed; 16 in each group. One-half of each group reproduced durations to its "memory" of the standard but once, after a delay of one or more days; one-half made one match to the standard after a two-minute delay, were re-exposed to the standard and made a second reproduction following a predetermined delay of one or more days. Random sequences were followed in determining the factors of: test order, delay interval and whether the subject made both a two-minute reproduction and another after prolonged delay, or a single reproduction only after prolonged delay. The delay periods between exposure to the standard and attempts to reproduce the standard by control of the variable stimulus were: two minutes and 1, 7, 14, 21 and 28 days (4, 5).

D. RESULTS

1. *Visual Duration (Flash Length)*

The duration of the standard stimulus and the mean reproduction time of the variable stimulus after delay are shown in Table 1 and have been graphed in Figure 1. The judged temporal equivalents fall consistently short of the duration of the standard, tending to approximate the actual interval as one of brief duration but demonstrating a pattern of regular underestimation.

TABLE 1
DURATION OF THE STANDARD STIMULUS AND MEAN REPRODUCTION TIME AFTER DELAY

| | Standard stimulus | 2 min. | 1 day | Variable stimulus | | | |
|--------------|-------------------|--------|-------|-------------------|---------|---------|---------|
| | | | | 7 days | 14 days | 21 days | 28 days |
| Flash Length | 2.00 | 1.59 | 1.43 | 1.81 | 1.57 | 1.54 | 1.21 |
| Tone Length | 1.50 | 1.02 | 1.44 | 1.53 | 1.36 | 1.32 | 0.69 |

There is a decrease in reproduction time, significant by *t* test ($P < .01$), for matches made at the same sitting, two minutes after exposure to the standard. Over the period of longer delays, 1-28 days, a similar reduction may be observed for reproductions made by subject control of the variable stimulus, a middle value of which (*M*) is also significantly shorter than the standard duration ($P < .01$), but this value does not differ reliably from reproductions made of the standard after only two minutes of delay.

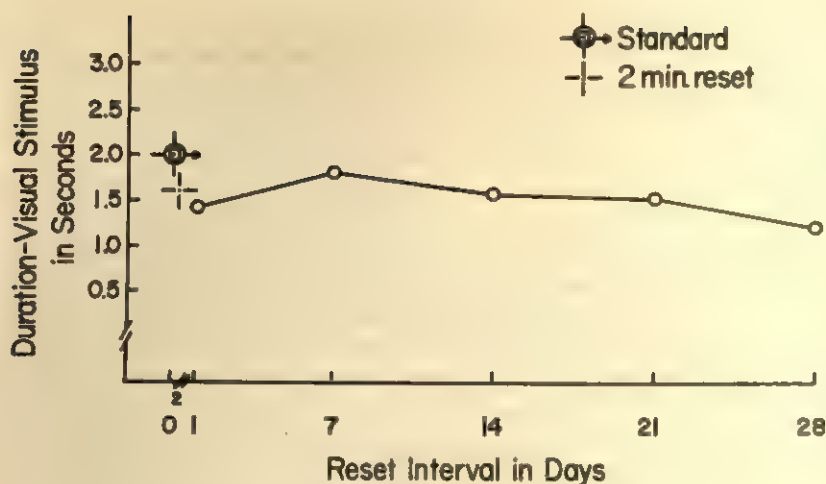


FIGURE 1

FLASH LENGTH OF THE STANDARD AND MEAN REPRODUCTION TIME AFTER DELAY

An analysis of variance was made to intercompare temporal reproductions made by the separate groups over the period of delay from 1-28 days and to contrast the performance of subjects making only a single reproduction with that of subjects making one reproduction at the same sitting and another after delays of one or more days. None of the F ratios computed was significant, either for main conditions or for interactions. The group variation remained relatively constant at all delay intervals.

2. Auditory Duration (Tone Length)

The duration of the standard stimulus tone and the mean reproduction time of the variable stimulus after all periods of delay are shown in Table 1 and Figure 2. Overall, the variable reproductions approximate the value of the standard stimulus, with a tendency to fall short of its actual value. There is a visible reduction in reproduction time, significant by t test ($P < .01$), for matches made at the same sitting, two minutes after exposure to the standard. Over the period of longer delays, 1-28 days, a similar reduction may be observed in the duration of the variable stimulus, a middle estimate of which (M) is also significantly below the standard duration ($P < .05$), but which falls significantly above the mean reproduction time observed after two minutes of delay ($P < .01$).

An analysis of variance intercompared performance by the subject groups over delay periods from 1-28 days and contrasted the performance of

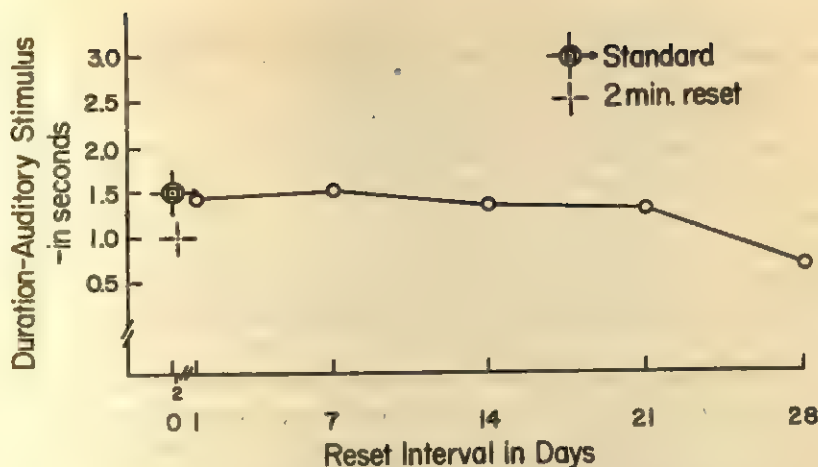


FIGURE 2

TONE LENGTH OF THE STANDARD AND MEAN REPRODUCTION TIME AFTER DELAY

subjects making only a single reproduction with that of subjects making one reproduction after two minutes of delay and another after delays of one or more days. None of the F ratios computed was significant, either for main conditions or for interactions. Group variance remained relatively constant at all delay intervals.

E. DISCUSSION

These observations, making use of Reproduction-Time as a measure of perceived duration, make it clear that naive human observers are able to approximate the duration of brief visual and auditory stimuli, whether the delay between registering the standard and matching its "trace" be as short as two minutes or as long as 28 days. The reproductions of standard stimuli fall short of actual duration in both modalities, but the error in performance was found to be no greater after prolonged delay (1-28 days) than had been noted after delays of only minutes and at the first experimental sitting. The tendency to underestimate the standard became conspicuous after the shortest delay, in keeping with what has been so often described in the literature on time-estimation for intervals falling in this general range (2, 7). No significant further underestimation was observed for interval reproductions made after longer delays, however, and there was, in fact, a significant improvement in accuracy for auditory duration for delays exceeding two minutes. There is also some indication of another

reduction in reproduction times after the longest delay period (28 days), but these differences are not statistically significant and may reflect only ordinary population differences among the subject groups.

An interesting question can be raised by contrasting the findings on the retaining of percepts of stimulus duration, as described here, and the observations made on the retention of standard stimulus intensity and frequency characteristics, that have been described in earlier reports (4, 5). The influence of the "direction of search" on judged perceptual equivalence has been commented upon (4, 5), meaning by this the tendency often seen for a point of judged equivalence to be displaced in the direction from which search was begun; as is commonly shown, for example, in determining sensory thresholds by the Method of Limits. The consequence of this tendency toward early-arrest as the point of equivalence is approached—in either direction—is usually taken into account in the balance of the experimental design or by some special form of statistical treatment of the data. Where the judgment is one of *duration*, however, the existence of such an effect would be in one direction only—as time moves forward—and it was this fact that determined the choice of the time-gating method of stimulus reproduction used in the present experiment. It might be expected, then, that the presence of a tendency toward early-arrest of the point of judged temporal equivalence would act only to reduce the estimates made of standard stimulus duration; a foreshortening, appearing in the data as interval underestimation. Although this methodological factor may not suffice to explain completely the characteristic underestimation of brief time intervals by the method of reproduction that have been so regularly recorded in the literature, and reflected in the data of this experiment, it is clearly a factor to be taken into account in any effort to explain the phenomenon.

No marked differences could be detected in the performance of subjects making but a single reproduction of the standard stimulus after prolonged delay and those making one reproduction at the same sitting and another after prolonged delay (1–28 days) for intervals presented by way of either sense modality. The Method of Reproduction had been selected for collecting data not only for its aptness and comparability to the kinds of non-verbal, manual adjustment of variable stimuli that had been used to observe the long-term retaining of other sensory characteristics (4, 5), but also for its ease, from the subjects point of view (1, 3), and its accuracy as a measure of the time-sense (6). The particular intervals singled out for use as standards were chosen because they fall in the range of the greatest possible accuracy in immediate reproduction (7), and yet give evidence of tapping

somewhat different levels of ability when reproduced after minimal delay in tests of time-judgment (1, 2). There probably exists a degree of difference in the stimulus strength of the standard visual and auditory sensory experiences to be carried through time which may account for the slightly superior performance observed in the reproducing of auditory as compared with visual stimulus durations in this experiment. There is rarely great exactness in the cross-modal equating of stimulus parameters and the values used in this study do not appear, subjectively, to be widely discrepant. A more explicitly designed test of the influence of this factor, however, remains to be explored in future experimentation.

The findings indicate, on the whole, that a human subject whose attention has been directed to the experience of a brief stimulus duration, delivered by way of the eye or the ear, can reproduce that interval with reasonable accuracy, characterized by a degree of underestimation. They further show that the error of underestimation does not increase for reproductions attempted after prolonged delay beyond that disclosed in performance under what would appear to be more optimal conditions, after only minutes of delay and at the same sitting. The reproduction of sample standard stimulus-durations thus showed little evidence of evanescence of the stimulus "trace" to which matches were made over the one-month period of observation. These data on the retaining of perceived duration, viewed together with what has been already described for the retention of the intensity and frequency characteristics of simplified percepts (4, 5), imply that "traces" of notable stability and longevity may be formed for sensory experience by the human subject, although experienced only briefly and in single exposure.

F. SUMMARY

The accuracy of human subjects in reproducing an experience of sensory duration following short and long delay periods was observed. Standard intervals presented via visual and auditory modalities were matched by a nonverbal reproduction after delays of minutes, days and weeks. The approximations made to a "trace" of stimulus duration showed remarkable stability, in both modalities, over all periods of delay, indicating little or no evanescence over the one-month period of study. A consistent underestimation was noted among interval reproductions of both visual and auditory stimuli. The greatest error in judged temporal equivalence was observed after the shortest delay.

REFERENCES

1. CLAUSEN, J. An evaluation of experimental methods of time judgment. *J. Exper. Psychol.*, 1950, **40**, 756-761.
2. FRAISSE, P. *Psychologie du Temps*. Paris: Presses Universitaires, 1957.
3. GILLILAND, A., & HUMPHREYS, D. Age, sex, method, and interval as a variable in time estimation. *J. Gen. Psychol.*, 1943, **63**, 123-130.
4. KING, H. E. The retention of sensory experience: I. Intensity. *J. of Psychol.*, 1963, **56**, 283-290.
5. ———. The retention of sensory experience: II: Frequency. *J. of Psychol.*, 1963, **56**, 291-298.
6. SPENCER, L. An experiment in time estimation using different interpolations. *Amer. J. Psychol.*, 1921, **32**, 557-562.
7. WOODROW, H. Time perception. Ch. 32 in *Handbook of Experimental Psychology*, Stevens, S., Ed. New York: Wiley, 1951.

Western Psychiatric Institute & Clinic
3811 O'Hara Street
Pittsburgh 13, Pennsylvania

A PREVENTIVE MENTAL HEALTH PROGRAM IN THE SCHOOL SETTING: DESCRIPTION AND EVALUATION*¹

*Department of Psychology, University of Rochester, and
Rochester (N.Y.) City School District*

EMORY L. COWEN, LOUIS D. IZZO, HAROLD MILES, EARL F. TELSCHOW,
MARY ANN TROST, AND MELVIN ZAX

I. INTRODUCTION

Quite beyond the geometrical burgeoning of technological complexity which has unquestionably characterized it, we have come to regard the 20th century, somewhat distinctively in contrast to its predecessors, as an epoch of increased preoccupation with, and focus upon, the psychological well being of man. Early and specific exemplars of this fact are to be found in: the emergence of psychoanalysis, not only as a method for the treatment of emotional disorders, but perhaps more important as a conceptual framework for the understanding of the dynamic aspects of human motivation and personality, and the development and subsequent growth of the mental hygiene move-

* Received in the Editorial Office on June 1, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ The authors would like to express their sincere appreciation to the following individuals for their cooperation in various phases of the project: Dr. Howard C. Seymour, Superintendent of Schools and Mr. Arnold B. Swift, Assistant Superintendent of schools in Charge of Administration, for their cooperation and foresight in helping to get the study off the ground; the late Dr. R. L. Springer who, subsequently, as Superintendent of Schools, offered encouragement and support in the continuing phases of the project; Miss Mildred Potter, Principal, No. 33 School, for her wholehearted interest and splendid cooperation in all phases of the data collection; Miss Mary Wintish and Mr. Ross Vasbinder, principals of the two control schools, who ably expedited data collection in those settings; Miss Helen Palmetter, Chief Consultant, Mental Health Clinical Services, Rochester Board of Education, for her continued encouragement, participation and helpful suggestions; Dr. Angelo Madonia for his continued support and valuable suggestions during the conception of the project; Dr. Herbert Greenberg and Mrs. Mary T. Kinsella of the Division of Parent Education and Child Development of the Rochester Board of Education, who contributed most meaningfully to the conduct and the evaluation of our parent seminar series; Mr. Joseph Hauser and Dr. Rita P. Underberg for their extensive participation in the process of data collection; Mrs. J. Frances Riley who as project secretary was a tower of strength during all phases of the total experience; Mrs. Sue Monjan, Mr. David Trost and Miss Anne Constantinople who labored effectively and diligently on major segments of the data analysis; and finally to Mr. Louis Goldberg of the New York State Department of Mental Hygiene who evidenced a sincere, wholehearted and constructive interest and participation in the project which far transcended the requirements of his position as Community Mental Health Representative for the N.Y. State Dept. of Mental Hygiene, Division of Community Mental Health Services.

ment. Experiences growing out of the international conflagrations of the present century have added considerable impetus to these earlier developments, by providing first hand experience for ever widening segments of our population with matters of emotional stability and its perturbations.

In our own country, a firmer foundation for mental health related activities was provided through the enactment of the National Health Act of 1946 allocating federal moneys for augmentation of existing facilities and services, extension of training, and acceleration of research. Support in each of these areas has increased steadily since the original legislation. Today the social sciences, biological sciences and medical sciences share a common interest in mental health matters and the knowledge and expertise of diverse professional specialists are, increasingly, being brought to bear on such problems. Doubtless, many would be willing to concede that a mental health orientation is firmly implanted as an integral element in the current American Zeitgeist.

Concern with mental health problems must, in the last analysis, be viewed as at least a double barreled phenomenon. There is the obvious and inescapable need to deal with identified and diagnosed instances of emotional disturbance by the most appropriate means available (i.e., the *curative* or *therapeutic* component) as well as the longer range need to reduce the source of such maldevelopment (the *preventive* or *prophylactic* component). The latter, because it usually tends to be less concrete and less pressing, has greater potential for being overlooked. Indeed it would not be unfair to say that the overall history of mental hygiene has been characterized by a domination of the therapeutic arm of the movement (i.e., the expansion of hospital and clinic installations, the training of more competent professional specialists, the extension of diagnostic and treatment facilities to schools, industry, the military, etc.). There can be little question about the worthwhileness of such a development *per se*. At the same time it must also be recognized that even a highly developed and effective therapeutic program is time consuming, extremely expensive, and restricted to a relatively limited segment of the population. Preventive work, by contrast, because it precedes the full blown occurrence of pathology carries with it the potential for reaching large segments of the population, early in their careers, and in a manner which vastly extends the effectiveness of resources of existing professional specialties. A strong case can therefore be made for the thesis that preventive approaches to mental health should be regarded as the primary focus of any comprehensive program designed to bring about a "saner society" of tomorrow.

For those who hold to such a viewpoint, it is a gratifying observation that

interest and research which begins to touch on aspects of psychological prevention, has made forward strides in the past decade to the point where such work is now a recognizable functional (rather than simply theoretical) component of the modern mental hygiene movement. Witness for example the publication by the Group for the Advancement of Psychiatry of a summary of the Bullis, Force, Ojemann and Forest Hills studies (8), the descriptions of the Harlem project (40), the Rockland County study (9), the massive and pioneering St. Louis Project (15, 16, 17, 18), the Michigan studies (13, 25, 30, 31), the work by the Bank St. College of Education group (1) and the California Study (3). Additionally, several volumes (5, 24) summarize recent conceptual and research advances in this field.

There are perhaps two "natural habitats" for research in the area of prevention—the home and the school. In our society these institutions share certain unique characteristics which suit them particularly well for the task. First and foremost, they are enduring and profound influences on the child during his formative years. His most significant identification models will come from the home and/or school situation, and the bulk of his learning and socialization derives from these sources. Certainly either of the two offers a good focus for research enterprises targeted to prevention. For practical reasons, such as geographical concentration, the availability of captive audiences, the likelihood of securing cooperation in experimental programs and assessment procedures, and the readiness for research that may have been engendered by ongoing contacts with mental health team members, the lion's share of preventive research to date has taken place in the school setting. Felix (14) writes spiritedly about the need for school participation in such endeavors.

Several unique attributes of research in psychological prevention warrant specific citation. For one thing, the term prevention itself implies the existence of some sort of program, which must extend over a period of time in its application. For this reason, as well as the fact that preventive programs are targeted to basic and complex psychological phenomena which may be expected to show change only gradually over long periods of time, it is to be expected that preventive research must be time consuming, and, in the last analysis, refractory to quick cross-sectional survey. Quite apart from the ultimately vital problems of follow up, even limited application and modest evaluation of a preventive program requires an essentially longitudinal approach to the research. A second natural attribute of a program in psychological prevention is that it requires the active participation of competent, well trained, professional specialists over substantial periods of

time. Hence from the standpoint of practicality, economics, and complexity, research in prevention is not easily done; quite likely such factors have been historical drawbacks to investigations in this area.

The project to be described in the present report is certainly a limited small-scale effort in any absolute sense of these terms. It has been confined, in essence, to a single school, has been able to draw upon the resources of only a limited number of professional specialists, and has utilized a rather narrow base of assessment procedures. Notwithstanding these failings the authors elected to undertake this project. The choice in these matters seems clear cut; one can either wait for many years until the "ideal" study can be carried out, or one can proceed as best possible within the realistic limitations imposed by the framework of available resources to provide leads and foci for future investigation, and to sharpen appraisal techniques. In this sense one can think of the analogy of a commando raid as opposed to a full scale invasion. We have elected to follow this "commando" alternative. Such pragmatism does not excuse the obligation to maintain a maximally realistic and self-critical attitude. It is our intention, insofar as is possible, to consider carefully the shortcomings of the present work later in this paper, as part of a final appraisal of its debits and credits.

In the sections which follow we shall seek to: (*a*) describe the preventive program which we have put together and applied in the school setting; (*b*) report on the instruments which we have developed and utilized in the assessment of the effectiveness of the program, the basic design of the research, and the procedures for data collection; (*c*) present a summary of our findings to date; and (*d*) discuss these findings in the context of presently available knowledge, in terms of the limitations of the research, and with respect to their implications for future work.

II. DESCRIPTION OF THE PROGRAM

A. BACKGROUND

For a city of its size Rochester, New York is, relatively speaking, well endowed with mental health services in the schools. Its 51 public elementary and high schools with approximately 42,000 students are served by a total of approximately 45 full-time psychologists and social workers. Although there are many different service patterns, depending on the size and location of the school, the average psychologist and social worker contact at any given school per week is 7/10ths time.

Such was roughly the case at No. 33 School, locus of the present research project, in 1957 when initial thinking about the research developed. At that

time the total clinical staff consisted of five half-days of social work time plus three half-days of psychologist time for a school with approximately 800 youngsters between kindergarten and 7th grade. The school neighborhood is best characterized socioeconomically as "upper-lower" and ethnically it reflects a fairly typical cross-section of Rochester.

The project evolved primarily because of the concern with the number of upper-grade level children being referred for behavior and personality problems or other indications of emotional disturbance. It was felt that too often children were not being referred until just before they were to leave for high school at which time their difficulties had already become well entrenched. The mental health clinical service (MHCS) workers, in sharing this concern with school personnel, found the latter willing to limit traditional services, so that available time and energy could be deflected to preventively oriented work in the primary grades. Early attempts in this direction included group psychological testing for first graders, brief "get-acquainted" interviews with mothers of first grade youngsters, and individual conferences concerning youngsters who had evidenced disturbance.

Fairly soon, however, it became clear that the program as conceived was too ambitious for the MHCS staff to do justice to on a part-time basis. In the spring of 1959 it was decided that in order to implement a more genuinely preventive program, the full-time services of a psychologist and a social worker for the primary grades would be required.² This was achieved shortly thereafter. Additionally, a part-time social worker and a part-time psychologist were assigned to the school to help meet continuing "traditional" service needs in the upper grades. The expansion of the preventive program also made possible a more systematic research evaluation. To facilitate both the clinical program and the research evaluation a consulting psychiatrist and a psychologist research consultant were added, to work cooperatively with the MHCS team. The program, as it has evolved, up to the present time is described below:

B. THE PRIMARY PROJECT

1. *Psychological Testing*

Except for a reading readiness test, administered in the first grade, children in the Rochester School System are not ordinarily tested until they

² The authors are grateful to the New York State Department of Mental Hygiene, Division of Community Mental Health Services for their financial support of the project during the three years from 1959 through 1962, as well as for support of this publication.

reach third grade. Thus, unless a child had been referred for individual study, the only available estimate of his ability and emotional status during his first three and a half school years was the opinion of his teacher. Our belief was that earlier and more detailed information might be of potential usefulness not only for optimal class placement but also to provide better awareness of the child's emotional status. Accordingly, a somewhat more comprehensive test battery was developed for administration at a very early stage in the child's school career. All first grade children are now tested in October, in groups of approximately 15. The testing is done in the classroom with the teacher acting as proctor and the psychologist administering the battery and acting as observer and recording behavioral responses during the testing situation. The test battery includes the following:

a. *An ability test.* The California Test of Mental Maturity, Short form (38).

b. *The Goodenough Draw-A-Man test* (19). Data derived from this test are used to supplement the California findings; additionally the psychologist undertakes a projective analysis of the productions.

To these test results, the psychologist adds his behavioral observations and the first grade teacher provides an ability estimate along a six-point rating scale. The composite of these data provides an early and preliminary assessment of the youngster's academic potential and emotional status. These findings are discussed routinely with the school social worker, prior to her interview with parents (*cf.* below). Similarly, wherever appropriate, the information is transmitted to pertinent school personnel (see section on "Teacher and Other Personnel Conferences").

2. Parent Interviews

In the earlier phases of the project the school social worker (SSW) held brief 20-minute individual interviews with the mothers of all first graders. In these interviews, information was gathered about the youngster's ability, personality, and behavior as well as the mother's interests and attitudes, at the earliest possible time. The assumption was made that in early detection, which was the aim of the interview, lay the best hope for dealing with evident or incipient difficulties before they had an opportunity to develop into major and enduring problems.

Because of the seeming importance of these early contacts, the project staff decided to extend the interview sessions to 55 minutes each and to exercise every effort to see mothers of all first grade children during the school year. In actual practice, either because of immutable work commit-

ments, or disinterest and lack of cooperation, it has been possible to see only about 80 per cent of the mothers each year. Characteristically mothers are given an initial appointment through a form letter signed by the SSW, psychologist, and principal, which indicates briefly the purposes of the interview. Every effort is made to provide a mutually convenient time for the mother. Up to two personal follow up letters are sent in instances where mothers have not responded to the initial communication.

Prior to the time the interview takes place the SSW has familiarized herself with the child's primary project folder which will, by then, include: identifying information, a description of the family constellation, comments by the kindergarten teacher, the results of standard reading readiness tests, health and attendance data, and the test findings and behavioral observations of the psychologist. To the extent that some unusual features are reflected in any of these sources, it is possible for the SSW to follow these up and to seek clarification during the contact with the mother.

The interviews are structured for the mother as part of a routine procedure for getting acquainted, in the context of the continuing interest of the MHCS team in the child and the family. Although the SSW may have some prior inkling of a problem situation, this is characteristically not discussed unless introduced by the mother. A modest percentage of the mothers (perhaps 10-15 per cent), who give evidence of the need to unburden or to ventilate or who directly ask for help, are scheduled for a limited number (usually no more than one or two) of additional interviews.

In broad scope the basic purposes of the interviews are as follows: (a) to establish a relationship with the mother; (b) to elicit pertinent background and current information regarding the child and his family; (c) to encourage mothers to express attitudes toward the school and toward education; (d) to explain both the educational objectives of the school and the role of the MHCS team; (e) to help mothers to feel free to contact the school or the MHCS team should they have concerns about the educational, emotional, or social well being of their child; and (f) ultimately, to formulate some tentative case-work impressions.

More specifically the SSW seeks to obtain information in each of the following areas: health and developmental history; attitudes toward eating, sleeping and toilet training as well as family practices in these matters; the pattern of interpersonal relationships in the family; the child's adjustment to playmates and to the larger community; the parents' perception of the child's functioning in, and adjustment to, the school situation; information concerning the parents' own education and present attitudes toward edu-

cation; data on the parents' employment, work history and socioeconomic status; and parental attitudes toward problems of sex and sex education, religion, parental roles, and major life goals.

General impressions of the SSW, as well as specific notations of unusual or ominous indicators are summarized in each child's Primary Project folder. Where incipient difficulties are judged to be fairly serious, the child's folder is clipped with a Red-Tag. To the extent that the mothers may have an awareness and an acceptance of such difficulties and a help-seeking orientation, referral procedures are initiated.

Primary Project folders are kept for all youngsters, whether or not the mother is actually interviewed. Where there has been no interview, anecdotal entries are made in this folder. Parent contacts subsequent to the initial interview are also recorded in the child's folder. When a child transfers into the school after first grade a folder is also prepared for him. Even though such folders lack the first grade testing and interview data, they will include materials from previous schools, teachers' comments, anecdotal observations and records of any contacts following the child's admission to No. 33 School. Hence some transfer children may also have Red-Tag folders, if they are known to be emotionally disturbed.

3. *Teacher and Other Personnel Conferences*

Following collation of the test and parent interview data, conferences are set up between the MHCS team and the first grade teacher, during which children representing the greatest source of concern are discussed. During the year, somewhat more than one third of the first grade youngsters are so considered. Pertinent data are shared and subsequent plans, goals, and objectives are formulated. An attempt is also made to establish an atmosphere in which the teacher feels free to recontact the team for additional assistance or to refer the child and/or his family for more intensive study and casework intervention. Additional conferences take place where necessary after the child leaves first grade.

It should be emphasized that neither the first grade nor the follow-up conferences are limited to teachers. Often they may be requested by the principal, the school nurse, the attendance teacher, the speech therapist, or special subject teachers. During the past two years the consulting psychiatrist has participated frequently in these conferences.

As a result of the combined effects of the testing program, parent interviews, and teacher and other personnel conferences, the number of requests

for intensive individual study of primary grade children has been reduced considerably in recent years. While we continue to receive appropriate referrals for intensive individual study, the overall reduction in the number of such requests is perhaps attributable to those aspects of our program described up to this point.

4. After-School Activity Groups

One of the unique features of the No. 33 Primary Project is the provision for two after-school activity groups, to which children giving evidence of special difficulties are assigned in terms of the best estimates as to their emotional and social needs. The diagnostic judgment concerning the child's need for such group activity represents the joint effort of the social worker, psychologist, and occasionally the consulting psychiatrist. It is based on the child's cumulative record in the school setting including pertinent data described in previous sections of this report, classroom observations by the teacher, and observations of other school personnel and the principal.

It is the purpose of these groups not only to provide a situation in which the child can have an opportunity to form meaningful interpersonal relationships in a structured but informal setting, but also to supplement the MHCS team's impressions about the child through reports about his reactions in a situation which is somewhat different from the everyday school setting. The groups are limited to 10 children each. The program, which is flexible, is planned jointly by the children and the group leaders in consultation with the MHCS team. Activities have ranged from woodworking to cookie-baking. The groups meet weekly after school for one hour on a specified day for 20 consecutive school weeks. Thus far, group membership has been confined to second and third grade youngsters known to the MHCS team through the Primary Project.

Group leaders have been teachers chosen from the No. 33 School faculty by the MHCS team. An attempt has been made to select leaders with a genuine and sincere interest in young children and their problems, an ability to function effectively with emotionally and/or socially deprived children, and finally, the attributes of being a good mother or father figure. Shortly before the group activity programs get underway leaders are provided with summaries of the children's backgrounds, and descriptions of their problems and needs as perceived by the MHCS team; in conference, limited goals are jointly established. Periodically, during the course of the activity groups' sessions, leaders meet with team members as well as with the psychiatric consultant to discuss their behavioral observations and impressions

and to take stock of progress to date. These conferences provide yet another source of information which may be used in planning for optimal placement in the school setting and for maximally effective methods of approaching the youngsters' emotional difficulties. Hence these after school activity groups are viewed as an essential facet of the overall No. 33 School Primary Project.

5. *The Consulting Psychiatrist*

Since early in the academic year 1959-60 a consulting psychiatrist has visited the school regularly for one half day each week and has participated in a variety of activities including the following: (a) conferences with the MHCS team about specific children; (b) conferences with the MHCS team, teachers, and other school personnel regarding individual pupils; (c) individual diagnostic work with children, on occasions, as indicated; (d) visiting classes to observe children in these settings; (e) participating in the promotion of mental health philosophy and principles in the school setting; and (f) participating in research and planning conferences.

6. *Parent and Teacher Meetings*

A series of separate programs for parents and for teachers was initiated during the academic year 1960-61. These were designed to promote a better understanding of mental health philosophy and facts, particularly as they pertain to areas such as personality development and functioning, human behavior, and the emotional and psychological needs of children.

Topics for the teacher seminar were selected by the MHCS team in consultation with a faculty committee. An original list of topics proposed by the MHCS team, was expanded by teacher suggestions. The teachers then reacted to the revised list, by indicating a preferential order of topics for consideration. On this basis a final slate was evolved. Five seminars conducted by the MHCS team members were held for the Primary Project teachers (kindergarten through third grade) centering around the areas of children's attitudes to the school situation, and indicants of anxiety in children's behavior. In addition to this the MHCS team invited a series of community specialists to meet with the entire No. 33 School faculty (including Primary Project teachers) to participate in a parallel series of six meetings. These latter meetings, some of which were also led by the MHCS personnel, included topics such as "The Socioeconomic Structure of No. 33 School District," "Problems Created by Differing Standards in the School and the Community," "The Application of Mental Hygiene Principles in the Class-

room," "The Effects of Emotional Deprivation," and "Mental Health for the Teachers." Thus, the Primary Project teachers were exposed to a total of 11 didactic-discussion seminars during the course of the academic year.

Parents of Primary Project children were invited, by mail, to attend a series of six meetings called "coffee hours." The purpose of these meetings was to provide a better understanding of the emotional and social development of children, as well as to provide an opportunity for group discussion in these areas. Meetings were set up at approximately two weeks intervals on the "buzz session" and discussion plan. The meetings were led both by a consultant of the Department of Parent Education of the Rochester Board of Education³ and the MHCS team members from No. 33 School. Each lasted approximately two and one half hours. The discussion topics included: "Discipline," "Social Relationships," "The Changing Roles of Fathers and Mothers," "Attitudes toward Responsibility and Authority" (two meetings), and "Sex Education." Attendance at these "coffee hours" was fairly good, ranging from a low of 24 on the evening of a very heavy snowstorm to a high of 70.

7. Planning and Research Meetings

Monthly planning meetings were held including MHCS team members, the principal, the school nurse (occasionally), the research psychologist, the consulting psychiatrist, the director and chief consultant of the Board of Education MHCS, the area representative of the New York State Department of Mental Hygiene (occasionally). At these meetings, consideration was given to the overall course of the project, progress to date, special problems, and plans for the future. Additionally, a subcommittee consisting of the Primary Project psychologist and social worker, the research psychologist, and the consulting psychiatrist met periodically to develop test instruments and assessment procedures, as well as an overall master design for collection of the evaluative data.

III. RESEARCH EVALUATION AND INSTRUMENTS

A. BACKGROUND

In its earliest stages the project was nourished and sustained by a belief, based primarily on observation and clinical impression, that its effects were largely salutary in nature. We have, however, been continually aware of

³ We should like particularly to express appreciation to Dr. Herbert Greenberg for his participation in this phase of the program.

the need to obtain consensual validation for these impressions by more objective assessment procedures. Accordingly, the first stage in a longer-range evaluation sequence was set up for the current academic year and has now been carried out. The present report will be confined to those findings emerging from this initial and preliminary cross-sectional appraisal. It is well to keep in mind that evaluation of a preventive health project must, in the last analysis, be evolved on a longitudinal basis; hence our present efforts are restricted to the establishment of "early trends" and initial points of comparison.

It is also appropriate to make explicit that objective assessment has been carried out within a framework of realistic limitations. Principal among these is the fact that the experimental (preventive health) program is confined to a single school and, although we have collected parallel data, wherever possible, in two demographically comparable control schools, the danger of confounding specific effects of the mental health program with other, not explicitly recognized attributes of the experimental school is ever present. In the ideal, the proper way to overcome such a defect would be to work with a sizeable sample of experimental schools. However, because of the extraordinarily heavy financial burden this would entail, it has been necessary to confine the program to a single school at this time. A second important reality limitation is the fact that parents and teachers have only a finite amount of time to devote to research assessment. Hence we have been limited in the depth and breadth of evaluation procedures which we could use. Our initial attempts at appraisal therefore suffer recognizably from their superficiality.

Two control schools were selected at the beginning of the experiment to parallel the experimental school as closely as possible in terms of the socioeconomic background and general intellectual level of their students. Such data were available from Board of Education records covering a number of years. For the preponderant majority of the measures to be described below, data comparable to those collected at the experimental school have also been obtained from these control schools.

B. ASSESSMENT PROCEDURES

Our various objective procedures, to be described below, fall into four categories. These are:

1. Pupil evaluation procedures.
2. Parent evaluation procedures.

3. Teacher evaluation procedures.
4. Prediction techniques.

1. *Pupil Evaluation Procedures*

Assessment in this area includes two classes of measures: (a) indices derived from school records and (b) test and appraisal techniques developed for purposes of evaluating project findings.

a. Indices derived from school records.

(1). *Referrals to the school nurse.* For each child in the school system an entry is made on the nurse's referral card each time the child is sent to the nurse's office for examination during the school day. Based on these records, two scores were derived: a *nurse's cumulative score*, summing the total number of such referrals during the child's school career, as well as a similar summary for the current (1960-61) academic year. The first of these two measures would obviously be incomplete for youngsters who transfer into the school system after the start of kindergarten.

(2). *Attendance record.* For the most part, records of the child's attendance for the academic year 1960-61 were available from the teachers' attendance cards kept at each school. In instances where this school attendance card was not available, attendance figures were obtained from the Teachers' Attendance Register filed with the Board of Education, the official record of the child's school attendance. Thus, for each child in both E and C schools an attendance score was derived, summing his total number of absences for the academic year 1960-61.

(3). *Grade point ratio.* Third grade children in the Rochester school system are marked on their report cards in seven different areas: Science, Social Growth, Work Habits, Reading, Language, Arithmetic and Social Studies. Any one of four grades may be assigned as follows: A (excellent), S (satisfactory), T (trying, but working below grade), and U (unsatisfactory). Based on the fifth and final grading period of the current academic year (1960-61), and using a code of $A = 5$, $S = 3$, $T = 1$, and $U = 0$, grade point ratios were determined for each child in the study. The highest possible score would thus be 35 (seven As) and the lowest zero (seven Us).

(4). *Objective test scores.* At the beginning of the academic year, all third grade children are given, in group administration, the Otis Quick-Scoring Mental Ability Test: Alpha Test, Form A (33). On the basis of this testing, verbal, nonverbal and total test IQ scores were available for all 8s in the E and C schools. Similarly, at the midyear all third grade children

take, in group administration, the S.R.A. Achievement Test: Reading, Form A, Grades 2-4 (41), yielding a reading comprehension and reading vocabulary score in percentiles. Hence, a total of five intelligence and reading aptitude scores was available for each child.

(5). *Achievement-aptitude discrepancy index.* Of some interest to us was the question of whether the child was performing up to potential. In order to quantify this variable, grade point ratios and total *IQ* scores for the combined E and C samples were converted into standard scores. Adding a constant of three to avoid negative scores, a D-score of these two standard scores was computed for each child, indicating the relationship between his grades (achievement) and his *IQ* (aptitude). These scores ranged from 0.67 at the lower end, reflecting marked underachievement to 6.78 at the upper end, reflecting marked overachievement.

b. Project test and appraisal techniques.

(1). *Mental health clinical services (MHCS) rating.* For each child in the E school, the psychologist and social worker independently went through the cumulative project record, and on the basis of it, made an objective rating on a five-point scale as follows:

(a). Know of no mental health difficulties. To designate children in this group (mostly transfers), who were given a "1" rating because they were not known to the team, the subscript "X" was appended to the rating. (There were 31 such "X" children.)

(b). Child has not given evidence of behavioral difficulty, but appears to have potential for this.

(c). Child has given at least some evidence of emotional difficulties.

(d). There is good likelihood of a serious problem here.

(e). Very definitely a seriously disturbed child.

The reliability of these independent sets of ratings was assessed using a Pearson Product Moment correlation. The resultant r of .929 indicated that the ratings were indeed highly reliable. Accordingly, for purposes of further data analysis the separate ratings were combined into a single pooled rating. The latter ranged from a low of 2 (no indication of any mental health problem) to a high of 10 (very definitely a seriously disturbed child).

(2). *Teachers' behavior rating scale.* An abbreviated version of a behavior rating scale used elsewhere (12) was developed for purposes of the present study. The scale included a listing of 17 characteristics (e.g., dependency, immaturity, destructiveness, disruptiveness, tendency to worry, moodiness, etc.) For each child in her class the teacher was asked to check

those characteristics which were considered to be applicable. For any characteristic so checked it was also necessary for the teacher to indicate the degree of severity of the problem on a three point scale: mildly, moderately, strongly. At the very end a final overall adjustment estimate was made, based on a five point scale ranging from very well adjusted (1) to very poorly adjusted (5).

Two scores were derived from this measure: a *total adjustment score* based on the number of items checked and their rated severity. If a child had been checked for all 17 items, and has been considered to manifest each one strongly, he would have had a maximum score of 51. The potential range then was from zero to 51. The second score, the *overall adjustment score*, is merely the rating on the last item, the range being from one to five. For each of these two indices, the higher the score the more maladjusted is the child considered to be.

(3). *Children's manifest anxiety scale (CMAS)*. The CMAS is a veteran measure, with an extensive background and history in the research literature (6, 7, 28, 29, 34), which was taken over without modification for purposes of the present study. Applicability of the measure to a third grade sample has been demonstrated elsewhere, and reliability data for this group have been provided (22, 23). In actual fact the CMAS consists of two subscales: a 42-item anxiety (A) scale, and an 11-item lie (L) scale. A-scale items have to do with symptoms or behaviors in the child, which would be indicative of manifest anxiety (e.g., "It is hard for me to keep my mind on anything," and "I get nervous when someone watches over me," etc.) L-scale items are designed to provide an estimate of the respondent's tendency to try to put himself in a particularly favorable light. This can be done either by claiming a degree of perfection that few people have (by answering "yes" to propositions such as "I like everyone I know" or "I am always good") or by denying typical motivations which may, however, have a socially disapproved component (e.g., "I would rather win than lose in a game").

The combined CMAS scale, following Holloway's procedure (22) was printed in oversized type, and distributed to each child in his own classroom setting. The teacher administered the scale following a standard set of instructions. The teacher read each item out loud, by number, and the child merely encircled the Yes or No response option, in accordance with what he considered to be applicable. The potential range of scores was from 0 to 42 on the A-scale and 0 to 11 on the L-scale, with higher scores indicating greater anxiety and greater lying respectively.

(4). *Secret stories.* The Secret Stories Test, adapted from the prior work of the St. Louis group (15), is a forced choice projective type test, consisting of three incomplete stories. The stories, which the child has before him in printed form, concern a little boy who is playing in the basement of his house only to be interrupted in turn by his mother, his father, and a friend. The administrator (in this case the teacher) reads the stories aloud to the child, and at the point of interruption, asks the child to close his eyes and to imagine how the story ends. On the basis of whatever ending he has conjured up, the child is then asked to select from a series of forced choice options first describing the child's feelings and then describing the child's actual behavior. Four adjectives—happy, mad, worried, and afraid—are paired in all combinations for a total of six forced choices per story. Similarly, four behavioral alternatives—"Says come play with me," "Hides behind the boxes," "Says go away" and "Does nothing and waits"—are also paired in all six possible combinations. Hence, for each child there are total scores for each of the four "feeling adjectives" and each of the four "behavioral options." Subscores for all eight of these responses are also available for "mother story," "father story" and "friend story," making a grand total of 32 scores per child. The Secret Stories Test was group administered by the classroom teacher and proved to be a motivating and rather enjoyable task for youngsters.

2. Parent Evaluation Procedures

Two basic instruments have been used in the evaluation of parents. The first of these is the Greenberg-Kinsella Scale of Attitudes toward Child Rearing (20). This scale, which has been used fairly extensively in evaluating aspects of the Parent Education program of the Rochester Board of Education consists of 31 items taken from already established scales having to do with various facets of child rearing (e.g., "If you ever give in to children, they will take advantage of you," and "Problems of children come out of troubles in the family"). The respondent is required to check one of four options: strongly agree, agree, disagree or strongly disagree. On the basis of these responses, scale scores presumably reflecting 10 child-rearing dimensions (uncertainty, strength of discipline, guilt, acceptance of negative feelings in the child, level of expectation, degree of permissiveness, willingness to allow the child to assume responsibility, comparison of self with with parents, comparison of self with child, and belief about extent to which child should learn from example of adults) as well as a total test score are obtained.

A second very basic parent measure used in the present study is an adaptation of the semantic differential (32). The research team developed a sizeable pool of potentially usable scales which were considered to be pertinent to our areas of interest (e.g., friendly-unfriendly; severe-lenient; interested-disinterested, etc.). From this pool the 16 judged to be most relevant were selected for the final instrument. Ten of these reflected the evaluative dimension and three each were selected from the potency and activity dimensions. The scales were set up in a random order of presentation on each page, and thus presented, the parents were asked to rate nine concepts: four buffers (Rochester, my neighborhood, newspapers, and television) and five critical ones (education, teachers, my child's school, my child's future, and mental health workers). The instrument was pilot tested on a group of 34 parents not associated with either the E or C schools in order to establish an optimal format and set of instructions for the major data collection.

The Greenberg-Kinsella was used on a pre-post basis to evaluate the parent seminar series, and both measures were used in the comparison of parent attitudes between E and C schools.

3. *Teacher Evaluation Procedures*

Two instruments were used for purposes of teacher evaluation. The first of these was a semantic differential test, very similar in format to the one described in the above section on parent evaluation procedures. The only differences were that two new scales were added making a total of 18, and there was some modification of the basic concepts to yield five buffers and six criticals. The latter included school teaching, parents, elementary schools, future, children, and mental health workers.

The second teacher measure was a modified Self-Ideal Sort, structurally similar to those used by Rogers and Dymond (36), Cowen *et al.* (12) and Cowen, Budin, Wolitzky and Stiller (10). A series of 28 items, each reflecting a potential role or function of the teacher, was identified (e.g., participation in school program, being in good physical condition, ability to help youngsters with personal problems) and a brief descriptive statement was drawn up for each. These items each reflected a positive teacher function. The rater's task was to order these along a seven point continuum, four at each point, ranging from, at one extreme, "best describes me," to "least well describes me" at the other extreme. This was done first from the frame of reference of the "ideal teacher" and second in terms of "myself as a teacher." Three scores are yielded by this sorting sequence: item placement for

ideal, for self, and the discrepancy (D-score) between these two placements. With 28 items, each having three such scores, a total of 84 such comparisons between E and C teachers were potentially available. Our prime concern is the D-score measure (which may also be summed, by teacher) since it has been proposed (35, 36) that the smaller such a score, the less the intra-individual tension and conflict may be presumed to be.

4. *Prediction Study*

In the final analysis, perhaps as pivotal as any aim of the present program was the wish to identify as early as possible those youngsters who are likely to prove to be poor mental health risks. To do so successfully and early in the game would enable a team of professional workers to focus their ameliorative and reconstructive efforts most efficiently. A unique opportunity for stock-piling data directed toward initial exploration of this problem was available within the framework of the experimental program. Each year at the experimental school, mothers come in during the spring semester to enroll their youngsters for kindergarten for the following year. We have specifically adapted two relatively brief screening instruments which, this year, we have administered to some 110 mothers during this registration period.

The first of these is a symptom checklist also adapted from the St. Louis project (15) listing a series of 20 items such as eating problems, sleeping difficulties, temper tantrums, etc. The parent is asked to indicate whether or not a given symptom is present and, if so, the extent of its severity on a four point scale ranging from very mild to very serious. The technique is designed not only to yield information on the incidence and severity of a wide variety of symptomatology in preschool children, but also is set up so as to provide a total maladjustment score.

The second screening device is a type of adjective check list in which the parent indicates the presence or absence of 34 personality or behavioral characteristics of the child (e.g., boastful, friendly, defiant, kind, respectful). For those characteristics selected as relevant, a two-point scale of applicability (somewhat, and very) is then applied. Once again data are available for each youngster, by individual item and in terms of a total scale score (i.e., how upset or maladjusted does the parent consider the child to be?).

These data will not be analyzed or discussed further in the present report. For certain of these youngsters, as has been true of past groups, we may expect that their project folders will be "Red-Tagged" at some time during the primary years when evidence of serious maladjustment or the potential

therefor develops from any one or combination of the sources which have become available to the MHCS team. Eventually we shall constitute a group of emotional and/or behavioral problems and compare the responses of the parents of such youngsters on the two prekindergarten scales, to those of parents of seemingly healthy children. Since this is a procedure which will be repeated annually it may be possible to cross-validate any empirically derived differentiating cluster. This particular phase of the total project has special appeal in its potential for early detection which many would regard a cornerstone in a comprehensive mental health program in the schools. At the same time it illustrates well the time-consuming longitudinal focus which contributes to making research in this field so demanding.

IV. METHODS AND PROCEDURES

A. SUBJECTS

In the course of the collection of data for the present study children, parents, and teachers, in varying numbers, were used as subjects.

1. *Children*

The experimental group (No. 33 School) consisted of all third grade children enrolled at the end of the current (1960-1961) academic year. For some measures (e.g., attendance records, grades, teachers' ratings) data are available for all 108 youngsters, comprising the four classes in this grade. On other measures, which required the physical presence of the child (e.g., Secret Stories, CMAS, Otis Intelligence), the *N* is reduced by the number of absentees on the particular date of test administration. In almost all instances such absenteeism did not deplete the total sample by a factor of more than 10 per cent.

The total figure includes a sizeable percent of children who transferred into the school at some time during or after first grade, as well as a much smaller percent of youngsters who had not started school with this group, but who had joined it somewhere along the way because of failure to have been promoted.

Control children were drawn from two ethnically and socioeconomically comparable schools (as indicated by census tract data), in close geographic proximity to the E school, totaling 97 and 79 third grade youngsters, respectively. There were also four third grade classes in each of these schools, and the percentage of transfers and failed promotions were comparable to those reported for the E school. For purposes of data analysis

the *N*s of the two control groups were pooled. Attrition in several measures due to absenteeism was at the same rate as that reported for the E school. The maximal total *N*, then, considering both E and C youngsters was 284 children.

Within the E school, several subsamples were drawn for purposes of special types of analysis. The first of these, known as the "Red-Tag" analyses, was based on youngsters whose project folders had been marked by the MHCS team at some time before the end of third grade to indicate the presence or likelihood of a moderate to severe emotional problem. Forty such youngsters, comprising 37 per cent of the total E school sample, were identified. These were contrasted with Non-Red-Tag children consisting of children who were perceived to be relatively or completely problem-free as well as children who had transferred into the school and were not yet well known to the MHCS team ($N = 68$ or 63 per cent of the grade).

In like manner, a subgroup analysis was done comparing children of parents who did and did not accept the invitation to come in for an interview with the school social worker while the child was in first grade. Because of the relatively high transfer rate, there were only 69 present third graders who could be considered as candidates for this analysis. Of these, there were 47 interview children and 22 noninterview children.

The foregoing figures for the "Red-Tag" and "interview" analyses are, of course, slightly reduced for those measures which could not have been taken by the child if he had been absent on the day they were given. Once again 10 per cent would be a good rough estimate of the maximal sample size reduction for this reason.

2. Teachers

A total of 18 E school teachers (nine primary and nine secondary) and 40 from the combined C schools (23 primary and 17 secondary) comprised the basic teacher samples for the present study.

3. Parents

For our major analysis, involving the semantic differential and the Greenberg-Kinsella child rearing attitude scale, there was a maximum of 73 usable records of third grade parents from the E school and 50 from the combined C schools. The relatively higher per cent of E school parents reflects the concerted effort which was made to secure the participation of E school parents (see section below on procedures).

In order for a parent to be included in the "pre-post" analysis on the

Greenberg-Kinsella measure, it was necessary to have been present at the first meeting in the seminar series when the pretest was administered, and to have attended at least two other meetings. These criteria, applied only to the E school where the parent discussions were held, yielded a "pre-post" sample of 22 including kindergarten through third grade parents.

B. PROCEDURES

1. *Children*

A goodly number of the children's measures required no special data collection procedures either for the E or combined C schools since they consisted of scores which could be taken from school and Board of Education files. Included in this category were the following: Nurses' Cumulative Referrals, 1960-61 referrals, Otis Verbal, Nonverbal and Total IQ scores, S.R.A. Reading, Vocabulary and Comprehension. Similarly the grade point ratios, MHCS clinical ratings and achievement-aptitude scores were determined at the conclusion of the 1960-61 academic year.

The teachers' ratings were done individually by each of the teachers, on her own time, sometime during the last month of the academic year. The CMAS "L" and "A" scales, and the Secret Stories were given in group administration in two settings by each of the E and C teachers, also during the final month of the year.

2. *Teachers*

Both of the teachers' instruments were group administered by a psychologist, not identified with the project, at a regularly scheduled faculty meeting of the respective E and C schools about one month before the end of the second semester.⁴

3. *Parents*

Pretest Greenberg-Kinsella attitude scales were group administered to parents attending the first of the seminar series, by psychologists not involved in the experimental program. These tests which were given before the start of the session itself, were structured as part of an independent outside research project. Posttests were also group administered at the end of the

⁴ In the case of the E school, an initial attempt at administration of the teachers' measures was thwarted by a misunderstanding concerning the nature of the task. Due to the lateness in the year and the lack of available time to schedule an additional faculty meeting, one of the members of the MHCS team administered the two instruments to the E school teachers shortly thereafter. The results on these measures should be evaluated with this bias in mind.

last or next to the last session in the parent series. In both the E and C schools, third grade parents were advised by letter that a special study was being conducted in which their cooperation was needed. In addition, the letters were buttressed in both E and C schools by telephone calls from "room mothers" to as many parents as possible seeking to enlist their help. Parents came in on a volunteer basis, in response to these requests, to a specifically scheduled meeting about a month before the end of the semester. At that time they completed the Greenberg-Kinsella and Semantic Differential Scales. Because of the wish to get as complete a sample as possible in the E school, after the regularly scheduled meeting, data were obtained from an additional sample of third grade parents on a home visit basis. Approximately 35 of the sample of 80 E school third grade parents were obtained in this manner.

C. HYPOTHESIS

It would be inappropriate to say that initiation of the research program was motivated by a series of finely drawn and sharply defined hypotheses. In point of actual fact, there was only one broad basic hypothesis underlying our work, and that was that a preventive mental health program in the schools would produce measurable positive effects on those people to whom it was targeted. The design of the research, and the very process of selection of assessment procedures was calculated to lend greater operational specificity to the test of this basic premise. Now, having described the mental health program, and the techniques by which it was to be evaluated, it is possible to articulate the implicit hypotheses in a somewhat more precise form.

First, however, it would be well to spotlight the fact that our investigation has been set up to tap three major problem areas:

1. Comparison of the E school (children, teachers and parents) with their respective C school peers.
2. Comparison of various subgroups with attributes relevant to the overall focus of the project within the E school.
3. Examination of relationships among variables without particular pertinence to this project, but nevertheless of significance to the larger corpus of psychological knowledge (e.g., what are the interrelationships among a variety of different types of adjustment measures in a third grade sample?).

Since these latter issues are varied and, in many instances, unrelated to central "project" concerns, such findings will be presented elsewhere in separate reports. Hence our development of hypotheses and presentation of findings will here be restricted to areas 1 and 2, above.

Although it should be relatively easy to perceive the common thread which runs through our enunciation of hypotheses, we shall present these separately for children, parents, and teachers.

1. *Comparison of E and C Groups*

a. Children.

(1). No differences between children of the E and C schools should be found on any of the three estimates of *IQ* derived from the Otis test. These are control measures which, along with census tract data and information about the socioeconomic and ethnic composition of the E and C schools, presumably demonstrate the initial similarity of the "treated" and "untreated" groups.

(2). The remaining seven school record measures, each of which reflects some component either of achievement or adaptation to the school situation, should favor children of the E school.

(3). On the basic children's adjustment measures—Teachers' Ratings ($N=2$), CMAS "L" and "A" ($N=2$), and Secret Stories, better adjustment should be found in the E children.

b. Teachers.

(1). Lower discrepancy scores between ratings of "myself as a teacher" and the "ideal" teacher should be found in E school teachers as contrasted with C school teachers.

(2). More favorable attitudes toward critical concepts (e.g., children, school teaching) should be expressed in the E group when compared to the C group on the semantic differential.

c. Parents.

(1). More favorable child rearing attitudes should be manifested on the Greenberg-Kinsella dimensions, by parents of E school children when compared to C parents.

(2). More positive attitudes toward critical concepts (e.g., my child's school, teachers, mental health workers) should be expressed by E parents in comparison to C parents, on the semantic differential.

2. *Comparisons of Subgroups Within the E School*

Two major subgroup analyses were carried out based on the children's data within the E school. The first of these compared the Red-Tag and

Non-Red-Tag groups, while the second compared children of parents who came in for an interview with those of parents who failed to come in for an interview. Additionally, wherever appropriate, comparisons were made of parents of these subgroups, as well.

Speaking now in terms of the Non-Red-Tag (nondisturbed) and interview subgroups in these separate analyses, we would expect: (a) better scores on the school record measures (e.g., grade point, reading vocabulary, etc.); and (b) scores indicative of better adjustment on the basic adjustment indices [see (a3) above]. In this case an additional measure, the MHCS clinical rating, applicable only to the E school, is also included.

Also pertaining only to the experimental school, we expect to find more favorable child rearing attitudes on the Greenberg-Kinsella scales in the posttest responses of those parents attending the parent seminar series.

Two final conceptual footnotes may be added. First, the analyses included in the present investigation, particularly those pertaining to children, were undertaken with some skepticism and reserve. As noted elsewhere, the process of change on basic personality dimensions is a slow and gradual one; hence it may be premature to expect to find clear-cut changes at this relatively early stage in the project. Particularly with respect to children, the need for long-range follow-up seems basic.

Secondly, although we feel that an explication of our guiding hypotheses is possible, it remains important to keep in mind that a considerable amount of data transcending the scope of these hypotheses, has been gathered. While we shall certainly be seeking answers to questions deriving directly from our hypothesis, we hope not to be constricted by these, and to utilize our broader findings as a source for the development of new hunches and leads to the maximal extent possible.

V. RESULTS

Data presentation will follow the same organization used in the previous section listing specific hypotheses. Comparisons between children, teachers and parents of the experimental and control schools will be followed by comparisons of subgroups within the experimental group.

A. COMPARISONS OF E AND C GROUPS

1. *Children*

School record measures of the children were taken from existing school files. Table 1 presents group means and variances for the basic control, school record and adjustment measures (excluding the CMAS). Total *N*s

for these comparisons range from 137 to 284, reflecting either absences on testing days or incompleteness of school records.

a. The three Otis *IQ* estimates represent control measures and were used to show that children from the E and C schools were comparable in basic potential. As anticipated, no significant or near significant differences are found on any of these measures.

TABLE 1
CONTROL, SCHOOL RECORD AND ADJUSTMENT MEASURES OF CHILDREN IN EXPERIMENTAL AND CONTROL SCHOOLS

| Measure | Mean | Variance | N | Mean | Variance | N | t | P |
|------------------------------|--------|----------|-----|--------|----------|-----|------|------|
| Control measures | | | | | | | | |
| Otis nonverbal <i>IQ</i> | 105.86 | 165.56 | 161 | 106.49 | 115.97 | 99 | 0.41 | N.S. |
| Otis verbal <i>IQ</i> | 101.96 | 164.18 | 161 | 101.61 | 119.14 | 99 | 0.23 | N.S. |
| Otis total <i>IQ</i> | 103.69 | 194.08 | 161 | 103.90 | 148.54 | 99 | 0.13 | N.S. |
| School record measures | | | | | | | | |
| Nurse ref., 3rd grade | 0.95 | 1.64 | 151 | 1.04 | 2.55 | 107 | 0.50 | N.S. |
| Cumulative nurse ref. | 2.87 | 5.24 | 47 | 4.14 | 18.60 | 90 | 1.90 | .06 |
| Days absent 1960-61 | 9.60 | 116.25 | 176 | 9.77 | 102.12 | 108 | 0.13 | N.S. |
| Grade point ratio | 21.46 | 20.27 | 176 | 21.69 | 28.78 | 108 | 0.38 | N.S. |
| SRA comprehension | 52.78 | 852.51 | 172 | 59.17 | 645.58 | 101 | 1.83 | .08 |
| SRA vocabulary | 53.57 | 896.67 | 171 | 56.88 | 713.02 | 100 | 0.91 | N.S. |
| Ach.-apt. discrep. | 2.97 | 0.98 | 161 | 2.98 | 1.06 | 99 | 0.50 | N.S. |
| Adjustment measures | | | | | | | | |
| Teachers' total adj. score | 9.56 | 74.28 | 175 | 10.03 | 74.72 | 108 | 0.42 | N.S. |
| Teachers' overall adj. score | 2.60 | 1.17 | 175 | 2.89 | 1.26 | 108 | 2.07 | .05 |

b. The remaining seven school record measures, each of which reflects either an achievement or adaptation to school index, fail to differentiate significantly between E and C children. Only two of these approached significance: cumulative nurses' referrals and SRA Comprehension. The former favors the control school ($p = .06$); it is, however, quite unreliable since data were available for only 25 per cent of the control Ss. In reading comprehension the E school children achieve higher scores, but not significantly so ($p = .08$).

c. Three instruments were used for the postexperimental comparison of E and C groups in terms of adjustive-behavioral measures. The first of these, the Teachers' Behavior Rating Scale, yielded both a *total adjustment score* and an *overall adjustment score*. On both of these measures (see Table 1) comparison between the E and C groups indicated directionally more mal-adjusted scores for the E Ss, this difference reaching significance at $p = .05$ for the *overall* score. This may mean that E school children were more mal-adjusted or that E school teachers making the ratings had become more

sensitive to maladaptive behaviors as a result of the experimental program for teachers. The basis for a choice between these alternatives will have to be provided by the results of the children's measures to be discussed below. Parenthetically, the correlation between these two teachers' rating measures for a total N of 283 was .67, indicating a substantial congruence for the two types of estimates. Separate correlations by schools were .76 for E and .62 for C, a difference which is significant when tested by a Z transformation. This suggests that E school teachers may have been operating from a better defined frame of reference toward emotional disturbance as the result of participating in the overall school program.

The CMAS yields two scores for each S , one on the Lie scale (L) in which higher scores reflect a stronger tendency to falsify and one on the Anxiety scale (A) in which higher scores reflect greater anxiety. The t -ratios arrived at in the several comparisons made on these data are reported in Table 2.

TABLE 2
GROUP COMPARISONS ON CMAS SCORES

| | Anxiety scale | | | | | Lie scale | | | | |
|----------------|---------------|----------|-----|------|-----|-----------|----------|-----|------|-----|
| | Mean | Variance | N | t | P | Mean | Variance | N | t | P |
| All boys | 19.73 | 69.79 | 144 | 2.64 | .01 | 4.22 | 7.30 | 144 | 2.01 | .05 |
| All girls | 22.46 | 74.05 | 127 | | | 4.80 | 4.21 | 127 | | |
| E school girls | 19.83 | 76.65 | 41 | 2.37 | .03 | 4.73 | 3.00 | 41 | 0.28 | NS |
| C school girls | 23.72 | 69.36 | 86 | | | 4.83 | 4.93 | 86 | | |
| E school boys | 19.48 | 72.36 | 61 | 0.31 | NS | 3.85 | 4.33 | 61 | 1.80 | .08 |
| C school boys | 19.92 | 66.41 | 83 | | | 4.48 | 4.62 | 83 | | |
| E school total | 19.62 | 76.37 | 102 | 2.06 | .05 | 4.21 | 3.93 | 102 | 1.75 | .09 |
| C school total | 21.85 | 69.30 | 169 | | | 4.66 | 4.74 | 169 | | |

The children in the E school average lower Anxiety and Lie scale scores than those in the C schools, indicating less anxiety and less tendency to falsify, but the t -test comparing them reached significance only in the case of the Anxiety scale score. Comparisons of Anxiety scale scores of boys and girls separately indicate that the differences between the girls in the two groups are significant and in the expected direction. When compared separately, E and C boys fail to show significant differences. It should be noted that

the correlation between Anxiety and Lie scale scores for the total sample ($N = 271$) was $-.10$ indicating that in this sample Anxiety scale scores are not contaminated by a falsification response set.

Differences between boys and girls as a group are significant on both the Anxiety and Lie scales with girls being more anxious and having a greater tendency to falsify. For the Lie scale only, the differences between the E and C boys tends to approach significance with the E boys showing less tendency to falsify.

The Secret Stories analysis was carried out by determining scores indicating the use of the four feeling responses (happy, mad, worried, and afraid), and the four behavior options (hides behind boxes, does nothing and waits, says "Come play with me," says "Go away"). Separate scores were available for the mother story, the father story, and the friend story and, by combining scores on these three stories, a total score was available. Hence, with eight response categories and four scores per category, a grand total of 32 scores per child was determined. E and C school children were compared by t -tests on each of these scores (the N s were 105 and 167 respectively). Of these 32 t -tests only two were significant with E children selecting the category "hides behind boxes" more and the category "says 'Go away'" less in connection with the father story. Since roughly two t -tests out of 32 might be expected to reach the 5 per cent probability level by chance, the observed pattern probably indicates that no overall significant differences exist between E and C groups on this measure.

2. *Parents*

Parents were evaluated by two instruments, the Greenberg-Kinsella Scale of Attitudes toward Child Rearing and the form of the Semantic Differential test developed by the research team.

In order to measure changes in parent attitudes on the subscales comprising the Greenberg-Kinsella measure, it was first necessary that the directionality of the so-called healthy responses be determined. Accordingly 10 psychologically trained judges were asked to respond to the 31 items making up the test under the set of trying to produce as healthy a record as possible. It was arbitrarily decided that if eight of the 10 judges could agree on direction of a given item, it would be keyed for its appropriate scale or scales. Of the 31 items in the total scale, 17 met the agreement criterion. As a result one factor was totally eliminated from the analyses and individual items were omitted from seven others. This left nine shortened subscales plus a total score comprising the sum of all items. In this pruned-down form posttreatment scores

of E parents were compared to the scores of C parents who took the scale only once. Of this total of 10 *ts*, one was significant at the .01 level, indicating that after the seminar E parents were more able to accept negative feelings in the child than were C parents.

On the semantic differential test nine concepts were rated on the same 16 scales. Since very few ratings by fathers were available, these were eliminated from the analyses. The elimination of a few more records because they were incomplete left the records of 61 mothers from E group and 51 from C group. Comparisons were made, using the *t*-test, between the ratings of the two groups on each scale used with each concept. The total number of *t*-tests was 144. The significant differences between groups are reported in Table 3.

TABLE 3

SCALES AND CONCEPTS ON WHICH SIGNIFICANT DIFFERENCES (AT THE .05 LEVEL) WERE FOUND BETWEEN EXPERIMENTAL ($N = 61$) AND CONTROL ($N = 50$) MOTHERS

| Concept | Scales |
|-----------------------|---|
| Rochester | calm-EXCITABLE*, FRIENDLY-unfriendly, humorous-SERIOUS, KIND-cruel |
| My neighborhood | calm-EXCITABLE, cooperative-UNCOOPERATIVE, good-BAD, pleasant-UNPLEASANT, strong-WEAK |
| Education | - - - - - |
| Television | - - - - - |
| Teachers | humorous-SERIOUS, LENIENT-severe |
| Newspapers | cooperative-UNCOOPERATIVE |
| My child's school | EASY-difficult |
| My child's future | INTERESTED-disinterested |
| Mental health workers | CALM-excitable, EASY-difficult, FRIENDLY-unfriendly, LENIENT-severe |

* The capitalized term denotes the direction in which experimentals exceeded controls.

Out of the 16 *t*-tests done per concept, roughly one might be expected on a chance basis to be significant at the 5 per cent level. Therefore, where there are only one or two scale differences on a concept, these may be regarded as probable chance phenomena. In Table 3 the concepts Rochester, My Neighborhood, and Mental Health Workers yield four, five and four significant scale differences respectively. The nature of these differences reflects a less positive view of neighborhood and a more positive view of mental health workers on the part of E mothers. The pattern of differences on Rochester seems to suggest that E mothers regard it somewhat more favorably than C mothers. The differences between groups on Mental Health Workers, one of the key concepts, were in the predicted direction and

are apparently not accounted for by a general response set to rate concepts more positively by the E group since, on one of the buffer items, My Neighborhood, the E group was less positive in its rating. There is some question as to whether these significant differences can confidently be regarded as reflecting genuine feelings, or whether they represent a need to respond in a socially desirable way. The fact that there were no significant differences in attitudes toward other key concepts would weaken the possibility of the latter. It should be noted that significant differences may occur between groups on semantic differential ratings without the occurrence of actual rating reversals between groups. Thus, in Table 3, the E group rated Mental Health Workers as significantly more calm than did the C group. This does not imply that C group saw Mental Health Workers as excitable; simply less calm. Actually there were no rating reversals on any of the scales showing significant differences in Table 3 or in any other similar tables to follow.

3. Teachers

Teachers in the E and C schools were compared on two measures. One was the Self-Ideal Q sort and the other was a semantic differential test constructed especially for their use.

The teachers' Self-Ideal Q sort consisted of 28 items, each describing a teacher role or function. These items were sorted twice by the teachers, once for "myself as teacher," and once for the "ideal teacher." Items were placed at any point from one to seven on a scale labelled "best describes" at the seven-end, and "least-well describes" at the one-end. Each teacher received three scores per item, one for the sorting for "myself as teacher," one for the sorting for the "ideal teacher," and one for the discrepancy between these two scores. Since the major orientation of the study is toward teachers of the primary grades, a set of 84 *t*-tests was done comparing differences in group means of the nine scorable records of primary grade teachers in E school and the 23 scorable records of teachers of the same level from the C schools. A total of six such differences were significant at the 5 per cent level or beyond. Since this total barely exceeds the chance expectancy of somewhat more than four significant differences, it was concluded that there were no differences between E and C groups on these measures. Carrying the analysis a bit further, a total discrepancy score summarizing all 28 individual discrepancy scores was determined and the means for the E and C group were found to be 35.67 and 39.69 respectively. Since the lower discrepancy score presumably indicates less conflict and greater satisfaction, these findings are in the appropriate direction. However, the *t*-ratio testing

the significance of the difference yielded a value of 0.85 which falls considerably short of a significant level. This analysis was repeated for the entire responding faculties of the E and C schools (N s of 18 and 40 respectively) and the means changed to 32.61 and 39.68, indicating a sharpening of the group differences of the prior analysis, in the expected direction. The resulting t testing the significance of these mean differences was 1.96. This is a value likely to be encountered by chance only six times in 100 and, while just barely short of the usually accepted significance level, suggests a tendency for lower discrepancy scores in the total E group than in the total C group.

The semantic differential test productions of Primary Grade Teachers, the main target group of the study, were analyzed separately from those of the Upper Grade Teachers. In this case 11 concepts were rated on 18 scales.

TABLE 4
SCALES AND CONCEPTS ON WHICH SIGNIFICANT DIFFERENCES WERE FOUND BETWEEN
EXPERIMENTAL ($N=9$) AND CONTROL PRIMARY GRADE TEACHERS ($N=23$)

| Concept | Scales |
|-----------------------|--|
| Rochester | ----- |
| Science | ----- |
| School teaching | friendly-UNFRIENDLY* |
| Literature | friendly-UNFRIENDLY, interested-DISINTERESTED, relaxed-TENSE |
| Parents | ----- |
| Newspapers | humorous-SERIOUS |
| Elementary schools | HELPFUL-harmful |
| Future | EASY-difficult, GOOD-bad, KIND-cruel, LENIENT-severe |
| Children | ----- |
| Art | cooperative-UNCOOPERATIVE, fair-UNFAIR, kind-CRUEL, lenient-SEVERE, optimistic-PESSIMISTIC, STRONG-weak, usual-UNUSUAL |
| Mental health workers | CALM-excitable, COOPERATIVE-uncooperative, EASY-difficult, FAIR-unfair, FRIENDLY-unfriendly, GOOD-bad, HUMOROUS-serious, INTERESTED-disinterested, KIND-cruel, OPTIMISTIC-pessimistic, PLEASANT-unpleasant |

* The capitalized term denotes the direction in which experimentals exceeded controls.

Table 4 indicates the scales and concepts on which significant differences were found between Primary Grade Teachers of the E and C group.

On four concepts, Literature, Future, Art, and Mental Health Workers, there are significant differences on three, four, seven, and 11 scales respectively. Of these four concepts, two, Literature and Art, were used as buffer items and toward both of these E group teachers seem more negatively disposed.

On the other hand, the other two significant concepts, Future and Mental Health Workers, were considered key ones at the outset and on both the E group teachers seem more positively disposed than C group teachers. The fact that, directionally, the two groups differ as they do on the four concepts suggests that no response set for seeing concepts favorably or unfavorably can account for these differences. Again the question arises as to the genuineness of the many more positive ratings of Mental Health Workers which were made by E group. In this case some partial control data are available which might help resolve the question of whether the data reflect a real feeling or a making of the rating that these teachers felt were expected of them by those under whose close scrutiny they found themselves. The ratings of the Upper Grade Teachers from E school represent such a control since these teachers, being in the E school, must have felt themselves under scrutiny but at the same time their actual contact with Mental Health Workers was more limited than that of the Primary Grade Teachers.

Comparisons between ratings of Upper Grade Teachers in E and C schools yielded more than two significant scale differences on only one concept, Science, a buffer. The E group viewed this concept as more cruel, more unfair, more pessimistic, but more helpful than did C group. On the concept Mental Health Workers, only one significant scale difference was found, a likely chance occurrence. This suggests that the opportunity afforded Primary Grade Teachers for much more than the usual amount of contact with Mental Health Workers, disposed them more positively toward such people.

B. COMPARISONS WITHIN THE E GROUP

1. *Children*

The group of children in the E school were subdivided in two different ways for the analyses to be reported in this section. The first of these contrasted children whose records had received Red-Tags for reasons described in the earlier section with all of those whose records had not been Red-Tagged. These two groups were compared by means of *t*-ratios on the 10 school record measures noted earlier as well as for five adjustment measures which have been described; the two Teachers' Ratings (total adjustment score and overall adjustment score), the two CMAS Scales and the MHCS rating. In Table 5 the group means, variances, *ts*, and probability levels are presented. In all cases *N*s are at or just under the total of 40 for the Red-Tag group and 68 for the Non-Red-Tag group.

We note initially that there are no differences between the groups on the

TABLE 5
CONTROL, SCHOOL RECORD, AND ADJUSTMENT MEASURES OF RED-TAG AND ALL NON-RED-TAG CHILDREN AND NON-RED-TAG CHILDREN EXCLUDING RECENT TRANSFERS

| Measures | Red-Tag | | All Non-Red-Tag | | Non-Red-Tag excluding transfers | | P |
|-------------------------------|---------|----------|-----------------|----------|------------------------------------|----------|--------------|
| | Mean | Variance | Mean | Variance | Mean | Variance | |
| Control measures | | | | | | | |
| Otis nonverbal <i>IQ</i> | 104.97 | 87.81 | 107.44 | 132.88 | 108.19 | 154.39 | 36 1.26 N.S. |
| Otis verbal <i>IQ</i> | 101.29 | 102.43 | 101.80 | 131.32 | 100.89 | 156.33 | 36 0.15 N.S. |
| Otis total <i>IQ</i> | 102.92 | 109.21 | 104.51 | 174.42 | 104.50 | 225.63 | 36 0.53 N.S. |
| School record measures | | | | | | | |
| Nurse ref. 3rd grade | 1.40 | 4.19 | 0.82 | 1.48 | 0.81 | 1.38 | 37 1.55 N.S. |
| Cum. nurse ref. | 4.11 | 13.28 | 4.16 | 22.29 | 5.46 | 26.70 | 37 1.27 N.S. |
| Days absent 1960-61 | 8.40 | 138.09 | 10.57 | 80.93 | 11.27 | 90.81 | 37 1.17 N.S. |
| Grade point ratio | 20.23 | 36.13 | 22.54 | 22.91 | 23.84 | 16.70 | 37 3.06 .001 |
| SRA comprehension | 52.03 | 584.24 | 63.66 | 641.24 | 65.63 | 570.18 | 36 2.45 .05 |
| SRA vocabulary | 51.84 | 776.92 | 60.09 | 657.42 | 60.71 | 707.49 | 35 1.43 N.S. |
| Ach.-apt. discrep. | 2.82 | 1.05 | 3.08 | 1.06 | 3.33 | 1.05 | 36 2.13 .05 |
| Adjustment measures | | | | | | | |
| Teachers' ratings | | | | | | | |
| (Sum) | 12.38 | 90.39 | 8.61 | 60.57 | 6.78 | 52.45 | 37 3.20 .01 |
| Teachers' ratings | | | | | | | |
| (Overall) | 3.25 | 1.22 | 2.69 | 1.17 | 2.14 | 1.84 | 37 3.96 .01 |
| CMAS (Lie scale) | 4.10 | 4.73 | 4.37 | 3.19 | 4.39 | 3.72 | 33 0.59 N.S. |
| CMAS (Anxiety scale) | 19.90 | 85.84 | 19.60 | 65.69 | 19.09 | 75.93 | 33 0.38 N.S. |
| MHCS rating | 6.25 | 3.63 | 2.44 | 1.27 | 2.76 | 2.02 | 37 8.95 .001 |

control, *IQ* measures. However, the Non-Red-Tag group scored significantly higher on Grade Point Ratio and SRA Comprehension. It also scored higher but not significantly so on SRA Vocabulary and lower but not significantly so on third grade referrals to the school nurse. Each of these differences is in the anticipated direction. Three of the five adjustment measures yielded statistically significant differences between groups and again each of these is in the predicted direction. MHCS ratings of the two groups were profoundly different. This, of course, was to be expected since the MHCS rating and the designation Red-Tag or Non-Red-Tag were highly interdependent! On the other hand the two sets of teachers' ratings, considerably more independent of the Red-Tag discrimination rating as indicated by correlations of only 0.25 between each of the ratings and MHCS rating, reliably discriminated between the two groups.

Since the group of Non-Red-Tag children which was used for the above comparisons included both children who were well known to the MHCS staff and 31 others who, by virtue of recent transfer into the E school, were not well known to the staff, the same analyses were repeated eliminating the scores of the 31 recent transfers. Hence while the *N* for the Red-Tag group remained constant, that for the Non-Red-Tag group was reduced by a maximum of 31. The results of these new analyses are also reported in Table 5.

The basic outcome of this analysis was similar to that reported for the entire Non-Red-Tag group with some sharpening of the findings. For example, three *t*-ratios which were significant in the expected direction in the prior analysis at $p = .05$ or $.02$ (the two teachers' ratings and grade point ratio) were significant at the $.01$ level in this analysis. Similarly, a new significant difference developed with Non-Red-Tag children displaying significantly higher achievement in relation to basic aptitude, when compared to Red-Tag children.

Responses to the Secret Stories Test were analyzed and the same 32 comparisons contrasting E and C group children were done between Red-Tag and Non-Red-Tag children in the E school. Only one of the 32 *t*-tests which were done was significant, indicating that this measure fails to discriminate between such groups.

A further analysis compared school record and adjustment measures of the children of parents who did and did not respond to at least two invitations to come in for an interview when the child was in the first grade. Since children who transferred into the E school late in the first year or sometime thereafter were excluded from this analysis, the *N*s were necessarily smaller. In the maximal case there are 47 for the Interview (I) group and 22

for the Noninterview (NI) group. The data for these comparisons are found in Table 6.

It is of some interest to observe apart from the three control variables that the results of 10 of the 12 comparisons in Table 6 are in the "expected" direction (i.e., better scores for I group children). These I group children

TABLE 6
CONTROL, SCHOOL RECORD AND ADJUSTMENT MEASURES OF INTERVIEW AND NONINTERVIEW GROUPS

| Measure | Interview | | | Noninterview | | | <i>t</i> | <i>P</i> |
|--------------------------------|-----------|----------|----------|--------------|----------|----------|----------|----------|
| | Mean | Variance | <i>N</i> | Mean | Variance | <i>N</i> | | |
| Control measures | | | | | | | | |
| Otis nonverbal <i>IQ</i> | 109.13 | 95.02 | 45 | 104.95 | 146.33 | 22 | 1.52 | N.S. |
| Otis verbal <i>IQ</i> | 104.58 | 93.39 | 45 | 94.95 | 146.24 | 22 | 3.53 | .01 |
| Otis total <i>IQ</i> | 107.56 | 129.21 | 45 | 98.55 | 195.88 | 22 | 2.82 | .01 |
| School record measures | | | | | | | | |
| Nurse ref., 3rd grade | 1.06 | 3.06 | 47 | 1.14 | 1.93 | 22 | 0.18 | N.S. |
| Cumulative nurse ref. | 4.72 | 24.20 | 47 | 5.55 | 11.78 | 22 | 0.70 | N.S. |
| Days absent 1960-61 | 9.26 | 78.24 | 47 | 13.00 | 216.57 | 22 | 1.31 | N.S. |
| Grade point ratio | 22.96 | 31.59 | 47 | 21.32 | 14.51 | 22 | 1.23 | N.S. |
| SRA comprehension | 61.11 | 546.92 | 46 | 56.14 | 699.23 | 21 | 0.78 | N.S. |
| SRA vocabulary | 63.85 | 699.95 | 46 | 47.30 | 536.43 | 20 | 2.42 | .02 |
| Ach.-apt. discrep. | 2.99 | 1.11 | 45 | 3.29 | 0.97 | 22 | 1.11 | N.S. |
| Adjustment measures | | | | | | | | |
| Teachers' ratings (Sum) | 8.64 | 67.54 | 47 | 8.55 | 56.83 | 22 | 0.04 | N.S. |
| Teachers' ratings (Overall) | 2.49 | 1.12 | 47 | 3.09 | 1.13 | 22 | 2.07 | .05 |
| CMAS (Lie scale) | 4.21 | 3.74 | 43 | 4.23 | 5.42 | 22 | 0.04 | N.S. |
| CMAS (Anxiety scale) | 16.88 | 79.44 | 43 | 21.95 | 64.52 | 22 | 2.24 | .05 |
| MHCS rating | 4.64 | 5.32 | 47 | 3.36 | 4.53 | 22 | 2.21 | .05 |

score significantly higher on SRA Vocabulary and show a tendency, though not a significant one, to have fewer absences and a higher Grade Point Ratio. In addition the I group children appear to be significantly more intelligent. In sum, the children of parents who responded to interview invitations appear to be more capable intellectually in terms of the foregoing school record indices. Several extremely interesting findings are to be noted in the adjustment measures. The teachers' overall adjustment rating discriminates reliably between the I and NI parents, with I children being seen as significantly better adjusted. Similarly, the CMAS anxiety scores of children of parents who were interviewed are significantly lower (less anxiety) than those parents who refused the interview. However, on the MHCS ratings

(which have a low positive correlation with Teachers' Ratings), I children score as significantly more maladjusted than the NI ones. This is a reversal, not only of prior expectancies, but runs contrary to the other two significant adjustment measure findings. Any attempt at explanation is highly *post hoc* and frankly speculative. Perhaps it is that failure to interview parents results in failing to uncover pathology that would otherwise have been detected.

A series of 32 *t*-tests was computed comparing the various Secret Stories scores of the children of the Interview and Noninterview groups. Of these, significant differences were found on only one of the analyses. This pattern of only one significant *t* out of 32 can be accounted for on a chance basis.

2. Parents

Preprogram Greenberg-Kinsella scores of E parents were compared with their own postprogram scores ($N = 22$) for each of the test scales. Of the total of 10 such *ts* that were done, three were significant at .05 level or beyond and one other approached significance, each in the predicted direction of "healthier" attitudes on the posttest. Specifically, it was found that parents tended to increase in their ability to accept negative feelings in the child ($p = .20$), increased in their wish to treat their own children differently ($p = .02$), increased in their wish to treat their own children differently from their own parents' treatment of themselves ($p = .02$), became less dogmatic about child-rearing practices ($p = .02$), and showed generally healthier child-rearing attitudes (Total Test Score, $p = .01$).

Parents of E school children were then divided up in two different ways and compared on their Greenberg-Kinsella scale scores. In both of these analyses which compared Red-Tag ($N = 29$) and Non-Red-Tag mothers ($N = 43$) and Interview ($N = 38$) and Noninterview mothers ($N = 11$) respectively, there were no significant or near significant differences. Thus despite the fact that this scale is apparently sensitive to changes in E parents due to the parent seminar, it failed to differentiate subgroups within the E group.

The same two parent subgroups were compared on their responses to the semantic differential test and the results of these analyses are found in Table 7.

Red-Tag mothers ($N = 24$) were found to differ significantly from Non-Red-Tag mothers ($N = 37$) on only a total of nine scales and in no case did the groups differ on more than two scales with reference to a particular concept. The findings are in no case much beyond what might have been expected by chance.

TABLE 7
 SCALES AND CONCEPTS ON WHICH SIGNIFICANT DIFFERENCES WERE FOUND BETWEEN INTERVIEW ($N = 36$) AND NONINTERVIEW
 ($N = 12$) MOTHERS, AND RED-TAG ($N = 24$) AND NON-RED-TAG ($N = 37$) MOTHERS

| Concept | Scales | |
|-----------------------|---|---|
| | Interview Noninterview | Red-Tag Non-Red-Tag |
| Rochester | *GOOD-bad, HELPFUL-harmful, INTERESTED-disin- terested, PLEASANT-unpleasant, STRONG-weak | RELAXED-tense, EASY-difficult |
| My neighborhood | IMPORTANT-unimportant, LENIENT-severe, PLEASANT-unpleasant, USUAL-unusual | KIND-cruel, HUMOROUS-serious |
| Education | ----- | ----- |
| Television | FAIR-unfair | calm-EXCITABLE |
| Teachers | GOOD-bad | ----- |
| Newspapers | FAIR-unfair, FRIENDLY-unfriendly, GOOD-bad, HELPFUL-harmful, PLEASANT-unpleasant, STRONG-weak | INTERESTED-disinterested, HELPFUL-harmful |
| My child's school | ----- | IMPORTANT-unimportant, cooperative-UNCOOP- ERATIVE |
| My child's future | FRIENDLY-unfriendly, HELPFUL-harmful | ----- |
| Mental health workers | HUMOROUS-serious | ----- |

* The capitalized term denotes the direction in which Noninterview and Red-Tag mothers exceeded the Interview and Non-Red-Tag mothers.

The Interview ($N=36$) and Noninterview ($N=12$) groups were found to differ significantly on more than two scales on three of the nine concepts that were rated. Most striking here is the fact that on every concept on which more than two significant scale differences were found, which includes no key concepts, the least cooperative group of parents, who refused to respond to several invitations to engage in an interview about their child's school adjustment, made more positive ratings than the Interview group. In fact, on all 20 of the scales showing significant differences between groups, the Noninterview parents exceeded the Interview parents in what may be considered the more positive direction.

This pattern found among Noninterview parents to excel in the positiveness of their ratings in direct contrast to what one might expect of such parents, suggests the operation of a response set. This set may well be related to a mechanism of denial which pervades their relationships. The manifest behavior of the mothers comprising the Noninterview group would support the possibility of the extensive use of such a defensive pattern. It should, however, be kept in mind that virtually every one of the foregoing significant differences occurs for a noncritical concept. Further research might profitably be directed toward elaborating or rejecting this possibility.

VI. DISCUSSION

At the outset it would seem proper to offer a concise interpretive summary of the extensive data reported in the previous section to highlight what seems to have been accomplished in this "commando foray."

A. COMPARISONS BETWEEN E AND C GROUPS: CHILDREN, PARENTS, AND TEACHERS

Comparisons of the E and C schools revealed no significant differences on the three *IQ* measures taken from the school files, indicating that, insofar as this type of variable is a basic one, the groups were reasonably comparable. In addition the schools were selected in terms of geographic proximity, and ethnic and socioeconomic similarity as judged by census tract data. The choice of the two control schools, on these bases thus seems to have been defensible.

The experimental program appears to have operated to make the E children less anxious as measured by the CMAS, this effect being attributable primarily to differences between E and C *girls*. Near significant differences between the Lie (L) scale scores of the total E and C groups, this time

due to lower L scale scores for E boys, suggest that the treatment program may also have operated to make E school children less defensive and more willing to admit negative symptoms or behaviors. The latter possibility adds additional lustre to the findings deriving from the A scale. The statistical independence of the A and L scales, as evidenced by an absence of correlation between them in this third grade sample, supports the belief that two separate factors have been touched by the experimental procedure. In each case the observed effects appear to be salutary ones, and are consistent with the objectives of the experimental team when the study was initiated.

These findings bring us face to face with a thorny, albeit classical, problem in the evaluation of relatively complex treatment procedures, not dissimilar to what the researcher studying the effects of educational procedures, drugs, or psychotherapy experiences in his efforts to assess a given agent of change. More specifically, the issue here is, what factor or factors in our fairly comprehensive and broadside approach to the creation of a school atmosphere which is optimally conducive to the promotion of good mental health, may underlie what appear to be reasonably stable, mutually supporting findings of some importance. Have we, by our parent interviews, encouraged a better acceptance of the school and its personnel which in turn operates to create a more relaxed atmosphere for the child at home? Have our parent "buzz sessions" helped to provide an outlet for a certain amount of anxiety in the parent thereby perceptibly, if minimally, reducing pressure on the child and leaving a somewhat less anxious and less defensive product? Is this change a function of the teacher discussion program, which has found a carry-over to the class room situation? Or is it that some of the more intensive aspects of the program such as individual psychiatric consultation, and the after-school groups, have worked to reduce anxiety in some of the more extreme cases? Clearly our data are insufficient to permit us to choose among such alternatives. This, however, should not obscure the importance of raising the issue, since in doing so we are implicitly pointing to a needed line of additional research. While it is permissible, perhaps even desirable, to start at a broad and comprehensive level in an experimental approach to prevention, if we are to be able to pinpoint cause and effect relationships, there is clearly an ultimate need to examine more microscopically the potential contributions of the various facets of the total program, as well as to focus on the processes that may be involved in change and not merely the final fact of change itself.

On the debit side of the ledger, the direction of the two teachers' rating measures indicated greater maladjustment for the E school children, significantly so in the case of the overall adjustment score. On the surface

the finding appears to contradict the preceding A scale and L scale effects. Here, however, there exists the possibility of a serious artifact, namely that in the process of orienting our teachers to problems of mental health and psychopathology, we may in fact have made keener diagnosticians of them. In so doing, we would have increased their sensitivity to the existence of pathology, thereby rendering them unequivalent to C school teachers as raters. The obvious control for this type of shortcoming insofar as future studies are concerned would be to develop a system of behavioral ratings whereby scores for all children are made by the same rater or different raters each of whom is independent of the program.

The Secret Stories Test failed to differentiate between E and C groups. This, however, seems primarily to be a function of certain structural limitations of the test, particularly the fact that in its present form there is an excess of "response pull" which operates to obscure individual differences. These limitations are discussed more extensively elsewhere (11).

The parents of E school children were found to have changed generally in their attitudes toward child rearing, to have become less dogmatic regarding such practices and to have increased in their wish to treat their own children differently from their own parents' treatment of themselves, after having taken part in the treatment program. As compared to C school parents, they were better able to accept negative feelings in their children. These changes are largely on a "within-school" basis, deriving from an analysis of differences of the pre-post test responses on the Parent Attitude Scale. As such, they in all likelihood reflect the effects of the coffee-hour, buzz-sessions series to which these parents were exposed. The opportunity provided to the parents in these sessions to explore, and to express feeling around such themes as attitudes to authority, discipline, and sex education was apparently sufficient to bring about constructive, positive changes in their publicly verbalized attitudes. This occurred notwithstanding some serious limitations in the criterion evaluation instrument, thus leaving open the possibility that a more effective and sensitive measure might have yielded a picture of even greater change. The verbalized reaction of the parents to the buzz-session series was a warm and acceptant one. Perhaps it reflected for them an interest on the part of the school and/or met a felt need to explore certain everyday, emotionally meaningful problems centering around their relationship to their children and their school. There are, of course, no guarantees concerning either the depth or the permanence of the observed changes. That they occurred however, is *per se* encouraging, and in the context of the apparently genuine acceptance which they were accorded, sug-

gests that the parent-seminar series may represent a prototype of a school-centered approach to prevention with considerable potential merit.

The *post hoc* shortening of the Greenberg-Kinsella instrument, which was necessary to develop a reliable keying of items for the psychological health of the various responses, unfortunately, as suggested above, limited the amount of useful information provided by the scale. This points to the necessity for replacing this instrument with another measure better suited to the purposes of the program.

The semantic differential test revealed differences between groups of mothers on only one key concept, mental health workers. It is felt that this probably reflects a genuine change in attitude in an area where attitudes were relatively amorphous and loosely held at best. We are, in our society, still in an era, where for large segments of the general public the terms social worker, psychologist, and psychiatrist, may conjure up vague and threatening images and associations such as "head shrinkers." One of the presumed virtues of the present program is the fact that many opportunities to modify such nebulous concepts are provided through varying contacts with members of the mental health team. Principal among these is the social work interview to which the preponderant majority of all first grade mothers have been exposed. Through this medium the parent comes to know *a* mental health worker in a very concrete form, and in a context of genuine interest on the part of the worker with an effort being made to establish the mental health team, and the school as interested friends. Quite probably the strongly positive perception of mental health workers by *E* school mothers reflects this concretization, and a diminution of the unknown mystical and potentially threatening aspects of professional mental health specialists. Since knowledge, and awareness of such specialists as potential sources of help for people in need of it is a prerequisite to seeking help, it is our belief that the significantly more positive attitudes of *E* school parents toward the concept Mental Health Worker, may be a change of more than minor importance. Attitudes toward the other key concepts—education, teacher, my child's school and my child's future—either were not changed by the program or else are so firmly fixed and stereotyped as not to show change on an instrument such as the Semantic Differential. Directionally the changes which were produced in *E* mothers were all as predicted. The failure to find more changes may be related to inherent limitations in the program, but it would also seem possible that a refinement of instruments is in order. For example, in line with some of the speculations about the semantic differential test results, it probably behooves the research team to make changes

in that instrument so that attitudes are given toward subtler concepts which should be affected by the treatment program.

The results of comparisons of the D scores of E and C school teachers' Self-Ideal Q sorts were in the predicted direction but fell short of usually accepted significance levels. Nevertheless, these differences, particularly those comparing the total participating faculties of E and C schools which were significant at the .06 level, are highly suggestive. At the very least these findings indicate that this instrument warrants further use.

Semantic differential test differences between primary grade teachers in E and C schools showed predicted changes on the concepts Future and Mental Health Workers. For upper grade teachers none of the predicted differences was found. It appears that the main target group among the teachers, those having primary grade classes, were more favorably disposed toward Mental Health Workers, as were E group mothers, after taking part in the treatment program. In addition they seemed more optimistic about the future.

The findings deriving from the two principal teacher evaluation measures are certainly not overpowering. Yet they are encouraging and, like the findings with the parents, typically in the anticipated direction. Apparently, some aspects of the total program have had salutary effects on the teachers' perception of their own role and function, as well as on their perception of the environment in which they work. The most specific of these changes again concerns the highly positive attitudes manifested toward mental health workers. One of the cornerstone features of the experimental program has been to provide the teachers with a genuine and meaningful on-going opportunity to get to know, to work with, to speak to, and to seek counsel from mental health specialists—the consulting psychiatrist, the school social worker and the school psychologist. This comes about in many specific ways: through conferences, formal and informal, about individual children, through meetings between the teachers and the consulting psychiatrist, through classroom observation and participation by members of the mental health team, each of which contributes to the development of an open, give-and-take atmosphere in the school setting. It is our belief that the mental health team has come to be perceived and accepted as an interested, helpful, and available resource in providing a more effective teaching situation, and that this is reflected in the significantly more positive attitudes by E school teachers toward the concept Mental Health Workers. The existence of positive, acceptant attitudes toward such specialists is viewed as an essential stepping

stone to the effective utilization of mental health resources in a preventive (or for that matter any type of psychologically oriented) program.

The foregoing factors may also contribute, at least indirectly, to the finding of a strong tendency toward less job tension, greater satisfaction with one's role as a teacher (Self-Ideal Discrepancy), and a more optimistic view of the future in the E school teachers. It may also be true that the existence of the teacher seminar series, using a minimal presentation of didactic material as a springboard to broad discussion of emotionally meaningful and central problem selections which are part of the teacher's everyday experience, contributed meaningfully to these self-ideal and semantic differential findings. In sum, each of these aspects of the program, communicates an interest in and a willingness to work cooperatively with the teacher in areas of vital everyday importance to her, to provide her with the feeling that there is support available in the inevitable moments of difficulty and uncertainty. Insofar as we can detect evidence of teacher change as a result of the experimental program, the straws in the wind appear to be encouraging and again point to the need and the worthwhileness of teacher-oriented features of a comprehensive mental health preventive program.

B. GROUP COMPARISONS WITHIN THE E SCHOOL

The comparisons of subgroups within the E group were certainly of equal, if not greater, interest than those between E and C groups. Of considerable significance is the fact that the behavior of Red-Tag children, about 85 per cent of whom were so designated by the school psychologist and social worker while in the first grade, was rated by their third grade teachers as significantly more maladaptive than was that of the Non-Red Tag group. This indicates that with the limited data garnered from test records and interviews with parents, a fair number of children with adjustment problems can be identified quite early, an essential factor if future preventive programs are to be optimally useful. In addition Red-Tag children were found to differ significantly in the predicted direction on their Grade Point Ratios, SRA Comprehension scores and in the Achievement-Aptitude Discrepancy scores as well as MHCS rating. The latter rating was made by the same psychologist and social worker who made the Red-Tag designation and it could not therefore be considered independent.

Several aspects of these findings with the "Red-Tag" group warrant further consideration. First, of course, the absolute incidence figures are, in and of themselves, striking. Given a more than usually close psychological scrutiny which was made possible by the concentration of psychological services

during the primary grades, 37 per cent of a class of third grade youngsters were identified, sometime between the ages of six to nine, as having emotional problems ranging from moderate to severe. It is instructive to compare these data with those from other incidence studies. For example, Bower (2) reviewed several studies which had been done in school settings and reported in one case that 42 per cent of the school children in Columbus, Ohio evidenced at least a moderate degree of poor adjustment. He reported also that in the Santa Barbara County, California schools it was found that from five to 35 per cent of the children, depending on the school district, were moderately or severely handicapped emotionally. Other incidence studies have been done with adults. Among these, Leighton's (26) finding that 37 per cent of those examined at random in a small town needed treatment corresponds most closely with the incidence figure reported herein. In their study of emotional disturbance among surgical patients Zwerling, Titchener, Gottschalk, Levine, Culbertson, Cohen and Silver (42) reported that nearly 55 per cent of their sample could be classified as neurotic or psychotic. Srole, Langner, Michael, Opler and Rennie (37), in the recently published Midtown study indicated that about 45 per cent of their extensive sample could be classed as being at least moderately impaired. Thus despite obvious discrepancies among the criteria utilized and the populations studied in these surveys, for the most part the incidence figure uncovered in the present study seems comparable to what is generally found.

From the important viewpoint of prevention, the incidence finding of this study seems most significant. The fact that such a substantial number of maladjusted children can be recognized as early as the first three grades has not been recognized up to now. Recognizing the practical import of early identification, Bower (2) demonstrated that teachers' ratings and certain standardized test scores could differentiate between emotionally handicapped and "normal" youngsters above the fourth grade. Present findings indicate that this can be done still earlier making it possible to optimally develop and direct preventive programs in the first few school years.

Our own "Red-Tag" children, so diagnosed by the mental health team, usually by the age of six or seven, already clearly show, as a group, certain unfavorable prognostic signs by the time they have reached the third grade. They are, given the same basic intellectual endowment, getting poorer grades than their peers, they are underachieving, and they are seen by their teachers as having a higher incidence of behavior problems. In the absolute sense, the incidence figures we have reported are discouraging and alarming. However,

if as the relationship between our own data and those of other incidence studies suggests, this is the nature of reality, it would seem to be exceedingly important to reach the earliest possible diagnosis of the situation on the presumption that it is more feasible to deal with such perturbations when the organism is still relatively young and pliable, and the pathology less entrenched. Hence one of the extremely useful potential by-products of comprehensive programs in prevention, is that of achieving the earliest possible diagnosis, and with it the application of treatment facilities where they are (a) most needed and (b) at a time when they are likely to be maximally parsimonious. To sum then, the specific procedure of "Red-Tagging," or more generally that of early diagnosis, should be a basic element of future preventive work.

Comparisons between the children of parents who cooperated in coming in to take the interview with the school social worker and those who did not come in likewise revealed several significant differences which were in the predicted direction. The finding that the Interview children were significantly higher on the Otis Verbal and the Otis Total *IQ* scores as well as the SRA Vocabulary scores suggests the possibility that parent interest in the project was related to their children's likelihood to do well at school because of native ability. On the other hand, it seems quite likely that the scores that such young children get on verbal tests depend heavily on the quantity and quality of verbal stimulation they encounter in the home. Therefore, one possibility for explaining such results is that parents who fail to come in for such interviews have difficulty communicating ideas and feelings generally and thus would avoid situations where they must engage in such behavior both in and out of the home. It is perhaps important to make such speculations and test them since the noncooperative parent will always pose a serious obstacle to the success of future preventive programs. If the reasons for such noncooperativeness can be fairly precisely identified, specialized techniques may be developed to stimulate the interest of these parents. The fact that Noninterview children were found to be more poorly adjusted by the Teachers' Overall Rating, and more anxious on the CMAS also supports the need to delve into the correlates of this avoidance behavior in parents. The one disconcerting element in this set of findings is the fact that the Interview children come out as significantly more maladjusted on the MHCS rating than do the Noninterview children. This finding is perhaps less alarming than it might seem at first blush, however, when we consider that the interview with the parent is a primary initial source for making inferences about maladjustive behavior and is also a primary datum for the MHCS

ratings. Hence, the absence of interview data with the mother could very easily account for a nonawareness of pathology and an assumption of its nonpresence in the Noninterview children, which would help to explain the "maverick" datum. The fact that both the teachers' ratings and the completely independent A scale scores go in exactly the opposite direction tends to support this line of reasoning. In any event, an attempt will be made to bypass this problem in future work, by making home visits to interview parents declining to come into the school for their interviews.

The parents' semantic differential test findings suggested the possibility that denial type defenses which limit communication characterize the Noninterview parents. Wherever significant differences were found between the Interview and Noninterview groups, the Noninterview group expressed the more favorable attitude. The consistency with which this was done and its contradiction of what one might expect suggests the operation of a response style for making uncritical, highly stereotyped responses. These findings seem to be in line with interpretations offered to explain Verbal Test Score differences between children of Interview and Noninterview parents. To this line of reasoning we can perhaps add one other observation. The preponderant majority of significant differences indicating more favorable attitudes by Noninterview mothers as compared to Interview mothers has occurred in the case of neutral concepts. If we are correct in assuming that the Noninterview mothers have manifested a defensive and stereotyped pattern of favorable responses, it becomes important to speculate why, in the face of such a response style, consistent differences did not also emerge on critical concepts as well. One explanation would be that these mothers have less favorable attitudes in the critical areas, but that their defensively positive responses were sufficiently strong so as to bring their scores to a par with Interview mothers for the critical concepts, at the same time exceeding the latter in favorability on the neutral ones. This line of reasoning, admittedly highly speculative, points again to the desirability of employing more subtle concepts in future work.

C. OVERVIEW

The study reported herein has combined the approach used in the pioneering Harlem Project (40) which focussed heavily on teacher training and the personal qualities of teachers which enhanced their ability to work with problem children and that of the St. Louis project (16) which concerned itself with the measurement and amelioration of parent attitudes. In providing skilled personnel with the full time responsibility for assisting in the

handling of mental health problems as they arose in the classroom, this study possessed some of the features of the program reported by Cutler (13). Of the many studies in this area which have been and are being done, this one in Rochester probably most resembled in its comprehensiveness the Forest Hills Project which was conducted in Toronto (8). This concerted attack from many angles has had its advantages in attempting to alter attitudes, or at least educate many of the people who played prominent roles in the lives of the youngsters who took part in the study, thereby enhancing the likelihood that the children were affected by the program. At the same time such an ambitious program suffered from certain weaknesses in its pristine form.

The major weaknesses of this particular project seem to exist in two distinct classes. Allusions have already been made to one of these; the need to develop instruments for measuring the precise changes in parents, teachers and children which could be reasonably expected to result from the program. This is a difficulty which has plagued many similar studies. Because of the unavailability of well-standardized and time tested measuring devices, there is relatively little overlap between the various devices used here and in other studies. The Secret Stories Test which was used in the St. Louis project to differentiate among children of different social classes failed to discriminate any groups in this study. The second class of weaknesses derives from the limited nature of the project. Generality of findings is diminished by the fact that the study focuses on a single experimental school from a certain type of community. As such we are subject to certain classic types of errors of design discussed elsewhere under the rubric of representative and systematic design in psychological research (4, 21). Thus, while the *Ns* in our E and C groups are stated as in the hundreds, in another sense our *Ns* are only *one* E school and *two* C schools. This of course increases the risk of confounding the presumably generalized effects of program with highly specific effects such as the personality of the school principals and their rapport with their staff, environmental conditions and facilities in a given school, and innumerable other such factors which are theoretically extraneous or orthogonal to the experimental program. There is, in the final analysis only one answer to this type of question and that is to buttress to the greatest extent possible the functional *N* on the critical dimension of schools, rather than on the immediately perceptible dimension of subjects. A principal drawback to such buttressing, already considered in an earlier section of this paper, is the extraordinarily heavy expense in time and budget that is a necessary part of the fabric of preventive research. Hence, to meet minimal

design requirements, the type of study reported herein must be supported by many independent replications suggesting that the research problem is one which can best be met on a national rather than a local basis. Alternatively, it will be necessary for communities to pool resources, if the present limited base is to be expanded meaningfully. In the view of the present authors, the stakes are sufficiently high to accord a high priority to either or both of the resolutions-in-principle.

The aforementioned limitations in the present study made it impossible to examine the relationships between social class and many of the variables that were studied, as was done in the St. Louis project (15). Another problem deriving from this source was the fact that the group of teachers in the experimental group, particularly those in the primary grades, was small and the not infrequent loss of a teacher from the program, due to a variety of causes, may have watered down the overall effects of the teachers' seminars. Increased scope longitudinally, in terms of follow-up work, is also called for. We are, after all, dealing, especially in our children, with very complex, basic and central processes, which may be expected to change only slowly over a very considerable period of time. Hence it seems particularly important to consider the present assessment to be the first in a series of evaluations designed to help us to separate transitory from enduring effects. Certain changes may be totally evanescent ones whereas others, perhaps not yet even seen, will accrete slowly over the years. As in psychotherapy, we are concerned in prevention programs not only with what may be observed in the final therapeutic hour, but also with how these changes endure, generalize and become modified with time. The ultimate fruits of this type of study are to be garnered in the form of a lowered incidence of emotional disturbance and a higher evidence of constructive school achievement. Such effects can only be viewed after a long period of time.

Currently plans are being made for a replication of the present study with modifications, hopefully improvements, in the parent and teacher programs and in the measuring devices that were used. Hopefully, the means will become available for expanding the project in order to include more of the schools in Rochester, N.Y., and perhaps schools in other communities of a different character. The encouraging results of this modest maiden effort seem to indicate a need for a more ambitious follow-up which can consolidate, elaborate and expand upon what has been found thus far. The premature termination of this type of project might well leave one with the same interesting leads, half-developed measuring instruments, and partly worked out parent-teacher programs which seem to typify many earlier studies.

The need for further elaboration of preventive programs in mental health cannot be overemphasized. The ever-expanding need, which seems never to be fully met, for mental health workers of all types clearly attests to this. So long as problems are permitted to develop in geometric proportions while capable workers are produced in arithmetic proportions, treatment personnel and facilities will be lacking. This seems to indicate that mental health workers, social workers, etc., who are in a position to deal with children before a fixed and often nearly immutable maladaptive pattern develops are more profitably used in a preventive program than in the more traditional role of diagnostician and therapist. This implies a readjustment in our thinking about the roles of such workers and their proper place in the school and the community.

There is perhaps some merit, at the very end, in releasing somewhat the normal constraints upon our imaginative processes, and in reflecting briefly toward the future. Through a shifting in orientation from diagnosis and treatment to prevention, from patch work and holding operations to an aggressive offensive, it may be hoped that the effectiveness of mental health specialists in the school can be increased manyfold so as to touch upon far greater segments of the population. This will not be the consequence of any single, limited research program, but rather, will come about from extended and concerted research efforts on a national basis, with vigor and imagination and the willingness to explore new hunches and approaches. We may hope that the preventive function will thereby be placed on an increasingly firm footing. We may even dare to envision several decades hence that school administrations and Boards of Education will be expending the bulk of their mental health budget on a before-the-fact rather than after-the-fact approach.

This much said . . . and "blue-skying" is always an enjoyable diversion . . . when we return to the level of fact and reality it is clear that the particular role the mental health specialist in the school might ultimately take, will depend on continued studies of the type described herein.

REFERENCES

1. BIBER, B. Integration of mental health principles in the school setting. In G. Caplan (Ed.), *Prevention of Mental Disorders in Children*. New York: Basic Books, 1961.
2. BOWER, E. M. Early Identification of Emotionally Handicapped Children in School. Springfield, Ill.: Charles C. Thomas, 1960.
3. ———. Primary prevention in a school setting. In G. Caplan (Ed.), *Prevention of Mental Disorders in Children*. New York: Basic Books, 1961.
4. BRUNSWIK, E. *Systematic and Representative Design of Psychological Experiments*. Berkeley, Calif.: Univ. California Press, 1949.

5. CAPLAN, G., Ed. *Prevention of Mental Disorders in Children*. New York: Basic Books, 1961.
6. CASTENEDA, A., McCANDLESS, B. R., & PALMERO, D. S. The children's form of the manifest anxiety scale. *Child Devel.*, 1956, **27**, 317-326.
7. CASTANEDA, A., PALERMO, D. S., & McCANDLESS, B. R. Complex learning and performance as a function of anxiety in children and task difficulty. *Child Devel.*, 1956, **27**, 327-332.
8. COMMITTEE ON PREVENTIVE PSYCHIATRY OF THE GROUP FOR THE ADVANCEMENT OF PSYCHIATRY. *Promotion of Mental Health in the Primary and Secondary Schools: An Evaluation of Four Projects*. Topeka, Kansas: Group for the Advancement of Psychiatry, 1951.
9. COOPERSTOCK, H. A mental health consultative service to public schools in Rockland County—A report on its first two years. Unpublished manuscript, Rockland County Mental Health Association, Inc., 1960.
10. COWEN, E. L., BUDIN, W., WOLITZKY, D. L., & STILLER, A. The social desirability of trait descriptive terms: A factor in the prediction of Q sort. *J. Personal.*, 1960, **28**, 530-544.
11. COWEN, E. L., IZZO, L. D., TROST, M. A., & MONJAN, S. V. The Secret Stories Test: A projective approach with young children. (In press.)
12. COWEN, E. L., UNDERBERG, R. P., VERRILLO, R. T., & BENHAM, F. G. *Adjustment to Visual Disability in Adolescence*. New York: American Foundation for the Blind, 1961.
13. CUTLER, R. L. A research evaluation of an action approach to school mental health: IV. Research evaluation of a school mental health program. *Amer. J. Orthopsychiat.*, 1961, **31**, 339-346.
14. FELIX, R. H. Mental disorders as a public health problem. *Amer. J. Psychiat.*, 1949, **106**, 401-406.
15. GILDEA, M. C.-L. *Community Mental Health*. Springfield, Ill.: Charles C Thomas, 1959.
16. GILDEA, M. C.-L., DOMKE, H. R., MENSCH, I. N., BUCHMUELLER, A. D., GLIDEWELL, J. C., & KANTOR, M. B. Community mental health research: Findings after three years. *Amer. J. Psychiat.*, 1958, **114**, 970-976.
17. GLIDEWELL, J. C., GILDEA, M. C.-L., & KANTOR, M. B. An evaluation of a preventive community mental health program: Progress report IV. Unpublished manuscript, Division of Research and Development, St. Louis County Health Department, 1960.
18. GLIDEWELL, J. C., MENSCH, I. N., DOMKE, H. R., GILDEA, M. C.-L., & BUCHMUELLER, A. D. Methods for community mental health research. *Amer. J. Orthopsychiat.*, 1957, **27**, 38-54.
19. GOODENOUGH, F. L. *Measurement of Intelligence by Drawings*. Yonkers-on-Hudson, N.Y.: World Book, 1926.
20. GREENBERG, H. M., & KINSELLA, M. T. The Rochester (ROC) parent attitude scale. Unpublished manuscript, City School District, Rochester, N.Y., 1960.
21. HAMMOND, K. W. Representative *vs.* systematic design in clinical psychology. *Psychol. Bull.*, 1954, **51**, 150-159.
22. HOLLOWAY, H. D. Reliability of the children's manifest anxiety scale at the rural third grade level. *J. Educ. Psychol.*, 1958, **4**, 193-196.
23. ———. Normative data on the children's manifest anxiety scale at the rural third grade level. *Child Devel.*, 1961, **32**, 129-134.
24. JAHODA, M. *Current Concepts of Positive Mental Health*. New York: Basic Books, 1958.
25. KIPFER, J. F. A research evaluation of an action approach to school mental health: I. Introduction to the action research project. *Amer. J. Orthopsychiat.*, 1961, **31**, 320-323.

26. LEIGHTON, D. C. Distribution of psychiatric symptoms in a small town. *Amer. J. Psychiat.*, 1956, **112**, 716-723.
27. LIPSITT, L. P. A self-concept scale for children and its relationship to the children's form of the manifest anxiety scale. *Child Devel.*, 1958, **29**, 463-472.
28. McCANDLESS, B. R., & CASTANEDA, A. Anxiety in children, school achievement and intelligence. *Child Devel.*, 1956, **27**, 378-382.
29. McCANDLESS, B. R., CASTANEDA, A., & PALERMO, D. S. Anxiety in children and social status. *Child Devel.*, 1956, **27**, 385-391.
30. McNEIL, E. B. A research evaluation of an action approach to school mental health: 3. The school mental health program. *Amer. J. Orthopsychiat.*, 1961, **31**, 332-338.
31. MORSE, W. C. A research evaluation of an action approach to school mental health: II. The mental hygiene dilemma in public education. *Amer. J. Orthopsychiat.*, 1961, **31**, 324-331.
32. OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The Measurement of Meaning*. Urbana: Univ. Illinois Press, 1957.
33. OTIS, A. S. *Otis Quick-Scoring Mental Ability Tests*. Yonkers-on-Hudson, N.Y.: World Book, 1939.
34. PALERMO, D. S., CASTANEDA, A., & McCANDLESS, B. R. The relationship of anxiety in children to performance in a complex learning task. *Child Devel.*, 1956, **27**, 333-337.
35. ROGERS, C. R. *Client-Centered Therapy*. Boston: Houghton Mifflin, 1951.
36. ROGERS, C. R., & DYMOND, R. F., Eds. *Psychotherapy and Personality Change*. Chicago: Univ. Chicago Press, 1954.
37. SROLE, L., LANGNER, T. S., MICHAEL, S. T., OPLER, M. K., & RENNIE, T. A. C. *Mental Health in the Metropolis: The Midtown Manhattan Study (Vol. I)*. New York: McGraw-Hill, 1962.
38. SULLIVAN, E. T., CLARK, W. W., & TIEGS, E. W. *California Short-Form Test of Mental Maturity*. Monterey, Calif.: California Test Bureau, 1957.
39. TELSCHOW, E. F., IZZO, L. D., MILES, H. C., TROST, M. A., & COWEN, E. L. Progress report of the preventive mental health project in a public elementary school, Rochester, N.Y. Paper read at the N.Y.S. Psychol. Assn. meeting, Rochester, N.Y., May 1961.
40. THE NEW YORK FOUNDATION AND THE BOARD OF EDUCATION OF THE CITY OF NEW YORK. *Report on the Harlem Project*. New York: The Board of Education of the City of New York, 1949.
41. THORPE, L. P., LEFEVER, D. W., & NASLUND, R. A. *SRA Achievement Series, Reading: Grades 2-4, Form A*. Chicago, Ill.: Science Research Associates, 1955.
42. ZWERLING, I., TITCHENER, J., GOTTSCHALK, L., LEVINE, M., CULBERTSON, W., COHEN, S. F., & SILVER, H. Personality disorders and the relationship of emotion to surgical illness in 200 surgical patients. *Amer. J. Psychiat.*, 1955, **112**, 270-277.

Department of Psychology
The University of Rochester
River Campus Station
Rochester 20, New York

CLINICAL JUDGMENT*

The American University and Sibley Memorial Hospital

DAVID J. KING AND RICHARD F. MANEGOLD¹

A. INTRODUCTION

This article is to report in the psychological literature the result of the application of a scaling technique to the study of physician's clinical judgments in two diseases. The particular clinical judgment concerned patient's prognoses. The detailed results of the study will be presented in the medical journal appropriate to the two diseases. The technique, however, would seem to be adaptable to many areas including clinical judgment in psychology.

In most reviews of the study of clinical and statistical prediction (1, 4) the conclusion is reached that, while both methods leave much to be desired, the statistical or actuarial approach is superior. Unfortunately, the empirical results have not seemed to inhibit some of the theoretical clinical writers who continue to wax poetic about clinical judgment (5, pp. 17-25).

Our present underlying thesis is that clinical measuring instruments (whether individual judgments or projective tests) have been found to be unreliable because of two basic considerations. First, clinicians have been required to make their judgments in terms of categorical alternatives rather than on a continuum. Thus, a predicting judge must state that a prisoner either will or will not commit a crime if released, or that a TAT protocol indicates that an individual falls into one of three or four categories. While there are some variations, nominal judgments seem to be the rule. A much more realistic procedure would be to require judgments on a continuum. For example, the judge might examine the MMPI protocols of 15 schizophrenics and place them in rank order with respect to favorability to reaction to psychotherapy. The second consideration that is suspected to produce unreliable results is that of the dimensions on which judgment is required. Clinicians are usually required to make judgments, from whatever data presented to them, directly to the actual event rather than on a more immediate psychological dimension. Judges might be asked if a subject will or

* Received in the Editorial Office on June 5, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ The work presented in this paper was supported by a grant from the National Institutes of Health, Grant No. RG-9812.

will not commit suicide, or if a patient will or will not recover from a heart attack. It is assumed that an easier judgment to make might be directly on the psychological dimension, asking such questions as, the probability of suicide or a recovery from a heart attack, or the two most likely and the two least likely alcoholics, out of 10, to stop drinking after a hospitalization.

The research reported herein has attempted to avoid both of these problems by requiring continuous judgments on an abstract psychological dimension.

B. METHOD

1. Subjects

The physicians who participated in this study were all practicing cardiologists or gastroenterologists (or specialists in internal medicine with a subspeciality in cardiology or gastroenterology). With one exception, all cardiologists utilized in the study were in practice in the greater metropolitan Washington, D.C. area (the one exception was in practice in New York City). The gastroenterologists were drawn from those practicing in Washington, Baltimore, Philadelphia, and New York City. The cardiological and gastroenterological groups were divided in an unsystematic manner into two groups, those to be used in the interview phase and those to be used in the testing phase. This division could not be done on a strictly random basis as the experimenters had to adjust their random selections to the restrictions imposed by vacations, time of availability, willingness to participate, etc. Cooperation was excellent. Of the 62 physicians contacted, all but three agreed to participate. Forty physicians (20 cardiologists and 20 gastroenterologists) were utilized in the testing phase, the remainder having been used as interviewees.

2. Procedure

The procedure will be presented, as it occurred, in three separate stages.

a. Interview. From a review of the journal literature and medical textbooks, a check list was devised for variables the investigators originally thought might be a prognostic importance in cases of Myocardial Infarction and L. Cirrhosis. Each interviewee was initially asked, "What information would you want to have to make an immediate prognosis in a case of Myocardial Infarction" (or L. Cirrhosis as appropriate)? It was explained that an immediate prognosis referred to the probability of the patient leaving a given hospitalization alive. As the interviewees responded, the items he mentioned that were on the list were checked while those items not on the

list were added. After this open-end stage, the interviewees were asked about the importance of the factors on the check list that they had not mentioned. In this report it would probably be unprofitable to go into detail about the information considered, but in general it included background information (age, sex, race, occupation, etc.), the nature and details about the current complaint, past medical history, physical examination, laboratory findings, treatment administered, and information concerning the training of the attending physician and the nature of the hospital (university or community) at which the patient might be hospitalized. Interviews were continued until no new information (or only obvious information such as that the infarcted patient was also not in the terminal stages of cancer) was elicited.

b. Patient history data gathering. Case histories of Myocardial Infarction and L. Cirrhosis were reviewed at the medical record rooms of two of the University Hospitals and two Community Hospitals. From the hundreds of histories of each disease reviewed, 20 of each were selected. The criterion for selection was based on the completeness of data. That is to say, those records were selected for use that came the closest to having the information that those interviewed had said that they wanted to have. As might be expected, not one record had all of the desired information. To keep the amount of information to be presented to the judges in the next phase, the information obtained was limited to that available from the first three days of hospitalization (and most of it from the first day). The information said to be of importance was transferred from the case histories to prepared eight-by-eleven-inch forms. One case history to a form. The final result of this work resulted in two sets of disease case histories, 20 Myocardial Infarction and 20 L. Cirrhosis.

c. Testing. In this phase the experimenters visited 20 cardiologists and 20 gastroenterologists and asked each to serve as a judge in establishing the prognostic scale values for the 20 patient histories appropriate to his specialty. Operationally, each judge was asked to place the 20 case histories in rank order in terms of favorability of immediate prognosis. An unlimited, but measured, time period was allowed for the judges to rank order the histories. Following the rankings, an interview was conducted and the judges were questioned concerning which variables they considered most important in influencing their judgments and what additional information they would want to aid their decisions. The results of these postjudgment interviews have no bearing on the present report and will be presented elsewhere.

C. RESULTS

The fundamental problem was with the determination of the consistency of the judgments of the two sets of physicians. Prognostic scale scores for each patient were computed by the normalized rank solution (2) for both diseases. The reliability of the scale scores were computed by a technique developed by Horst (3). The reliability of both sets of scale scores was .97. The authors suggest that the ranking procedure is a powerful technique for the investigation of the consistency of clinical judgment.

It is also possible to obtain an indication of the validity of the judgments by comparing the scale scores on the patients with those who actually died during that hospitalization. Three of the L. Cirrhosis patients and four of Myocardial Infarction patients expired. Point biserial correlations were computed between the prognostic scale scores and the living-expiring dichotomy. The correlations were .43 and .61 for the L. Cirrhosis and Myocardial Infarction cases respectively. The .05 and .01 levels of significance are .44 and .56, respectively. When it is considered that the data presented to the judges were limited to that obtainable during the first three days of hospitalization only, and that some of these patients expired after weeks in the hospital, the above results are most encouraging.

In neither disease did the time spent by the judges in making the judgment have any effect on the validity of the judgments. For the cardiologists working with the Myocardial Infarction histories, there was also no influence in terms of the reliability of the judgments. For the gastroenterologists, however, the time taken in making the rankings seemed to influence reliability. Reliabilities were separately computed for those 10 judges taking the longest time to rank and for those taking the shortest time to do the ranking. The reliability of the scale scores for those 10 judges taking the longest time was .92, while it was .97 for those 10 judges taking the shortest time.

D. DISCUSSION

This technique of research has several possible applications. While the present discussion will deal with situations in the medical field, there are many direct extrapolations to problems in psychology.

The first obvious extension of this research would be in the study of the development of the consistency of clinical judgment. It will be interesting, for example, to compare the reliability of the judgments of the tested specialists with those of general practitioners, residents at various levels of

training, interns, and medical students through each year of training. One would predict that a medical student just starting training would rank the case histories in an inconsistent manner. As training and experience increased, one would expect the reliability of judgments to increase. The pattern of growth of the consistency of clinical judgment could also be traced. Does clinical judgment increase regularly or in spurts? How much consistency of clinical judgment is obtained from classroom and textbook experience and how much from experience with patients?

Another interesting question concerns the degree of generalization of clinical judgment from one area of specialty to another. Thus, the cardiologist would be asked to rank the histories of L. Cirrhosis and the gastroenterologists those of Myocardial Infarction with the results of both compared with the reliability of judgments of general practitioners. Two different predictions seem reasonable. One point of view is that specialization and continued study sensitizes and increases clinical judgment which will transfer from one disease to another (note that the diagnosis is given to the judge and is not part of the test). Another suggestion is that increased specialization narrows one's reading, experiences, interests, etc., and that judgments remain specific to a given area of study.

With more adequate data, it would be possible to extend the statistical analysis of the results to determine the relative importance of the various pieces of information in influencing their judgments. Thus with a larger number of cases and with all data complete (i.e., no missing values) as might be possible in a prospective rather than a retrospective study, it would be possible to predict the scale prognostic scores through a multiple regression analysis. Examination of the Beta weights of each variable in the regression equation would indicate their relative importance in determining the scale scores.

The multiple regression analysis also has the possibility of finding those variables with a zero or insignificant Beta weight and, if this information is not needed for other reasons, they might not be routinely obtained. This would conserve the time of the medical personnel and save the patient money.

E. SUMMARY

Twenty case histories of Myocardial Infarction and L. Cirrhosis were presented to specialists in each of these diseases. The specialists were asked to rank order the case histories (appropriate to their specialty) in terms of favorability of immediate prognosis. Prognosis scale scores were computed

for each disease and their reliability and validity tested. The reliability was .97 for both scales and the point biserial validities were .43 and .61 for L. Cirrhosis and Myocardial Infarction, respectively. Implications for this research were discussed.

REFERENCES

1. GOUGH, H. G. Clinical versus statistical prediction in psychology. In *Psychology in the Making*, Postman, L., Ed. New York: Knopf, 1962. Ch. 9.
2. GUILFORD, J. P. *Psychometric Methods* (2nd ed.) New York: McGraw-Hill, 1954. Ch. 8.
3. HORST, P. A generalized expression for the reliability of measure. *Psychometrika*, 1949, **14**, 21-31.
4. MEEHL, P. E. *Clinical vs. Statistical Prediction*. Minneapolis: Univ. Minnesota Press, 1954.
5. THORNE, F. C. *Clinical Judgment*. Brandon, Vermont: Journal of Clinical Psychology, 1961.

Department of Psychology
Albion College
Albion, Michigan

LYSERGIC ACID DIETHYLAMIDE (LSD-25): XXXIV. COMPARISON WITH EFFECT OF PSILOCYBIN ON THE SIAMESE FIGHTING FISH*

South Oaks Research Foundation, Inc., Amityville, New York

H. A. ABRAMSON, H. H. GETTNER, A. ROLO, AND G. DEAN¹

A. INTRODUCTION

Young adult and juvenile forms of the Siamese fighting fish have been used to assay LSD-25 and related compounds (1, 2, 3). The bioassay method previously published has been improved, and a statistical procedure suitable for comparison of LSD-25 with other compounds has been evolved. The vegetative, motor and behavioral responses of the Siamese fighting fish swimming in solutions of LSD-25 in concentrations as low as 0.1 mcg per ml are characteristic. Briefly, after a period of excitation, fish responding to the LSD-25 become semiquiescent and assume a characteristic position at the surface of the liquid. The position is nose up—tail down, with the body of the fish at an angle of about 40° to the surface of the liquid. Exaggerated backward movements, Cartesian diver effects, peculiar rolling of the fish on its long axis, kinking in body formation, pigmentation effects and display of the fins are usually observed. It is the purpose of this communication to compare the effects on fish exposed to LSD-25 in the outside liquid with those fish injected with LSD-25 or with Psilocybin. It has been found that the reaction-dose curves are essentially parallel. It seems likely, therefore, that LSD-25 and Psilocybin act on the same enzyme system as measured by the reactions observed.

B. METHOD

The nose up—tail down and Cartesian diver positions have been used for statistical evaluation of the response. After treatment with LSD-25, the fish slowly return to normal. The length of time required for recovery depends on the size of the dose and the method of administration. It may be

* Received in the Editorial Office on June 6, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ From the South Oaks Research Foundation, Inc., Amityville, New York. This research was begun at the Biological Laboratory, Cold Spring Harbor, New York and is being supported, in part, by a grant from the Sport Fishing Institute, Washington, D.C.

mentioned that responses somewhat similar to LSD-25 are obtained by exposing the fish to derivatives of LSD-25 like 2-brom-d-lysergic acid diethylamide (BOL), as well as d-lysergic acid ethylamide (LAE). The bioassay procedure was carried out in a room devoted solely to maintaining populations of Siamese fighting fish of two varieties, *B. splendens* and *B. cambodia*. The fish were kept in distilled water treated with salts (NaCl 0.176 gm/l + KCL 0.086 gm/l), since it was found that reactions to drugs of fish not conditioned by exposure to salts were apt to be variable under our experimental conditions. Fish considered young adults are four cm or more in length. The juveniles are shorter. Ten fish were observed simultaneously in square glass jars, in 100 to 150 ml of liquid. Observations were made every five or ten minutes, depending upon the nature of the experiment. The room was kept illuminated at all times. The temperature of the aquaria varied between 78° and 80° F. During the summer, higher room temperatures may prevent experiments in this range. All the aquaria were aerated and equipped with filters. The fish were fed a uniform diet twice a week.

Considerable flexibility in the experimental procedure has been achieved by developing a new technique for injecting small volumes of liquid, without anesthesia, directly into the body cavity of fish weighing as little as one gram. Previously, experiments were essentially restricted by the need to place the fish in a comparatively large volume of solution (at least 100 ml) containing a given quantity of the material under investigation. Because of the large volume of solution, experiments were necessarily limited by two factors: (a) the scarcity of the two compounds to be studied; and (b) the gills had to be permeable to the compounds. With our present technique, using juvenile or young adult forms of the Siamese fighting fish, only 0.05 ml is required for each injection, thus eliminating both of the foregoing limitations.

The fish were injected through the tail into the body cavity directly, without anesthesia. Since the fish are air breathers, they can be kept outside of water for several minutes without injury. They were immobilized by cellulose acetate sheets with the injection made with or without a binocular microscope. Thirty-gauge needles were employed. After injection the fish were observed in square jars 16 cm high, containing 300 ml of distilled water and kept in a water bath at 78° F.

In experiments with LSD-25 and its derivatives, as well as with other compounds that do not kill the fish, the fish are kept for a recovery period of three weeks before being used again.

Observations were often made by two individuals, one of the observers

being unaware of the nature of the experiment. To offset the presence of bias when experiments were not run blind, photographs with a Polaroid camera reading were used to check the data.

A wholesale house supplied the laboratory each week with large quantities of healthy young adult and juvenile fish, conditioned by establishing equilibrium with salts in the aquarium water as mentioned hitherto.

C. RESULTS

1. Part I: LSD-25

Data obtained previously were checked by making a study of 290 Bettas exposed to varying concentrations of LSD-25 in the outside liquid.² Figure 1 shows the averaged per cent of fish reacting as a function of time. The concentrations studied were 0.1 mcg per ml, 0.5 mcg per ml and 1.0 mcg per ml. Ten fish per jar were observed, the experiments in all cases lasting four hours, with five experiments of 10 fish each used as controls. In order to obtain regression lines, three lines were computed at mean times of 1-1/2 hours, 2-1/2 hours and 3-1/2 hours. A mean time of 1-1/2 hours consisted of readings of counts of y , the per cent reacting at 1:20, 1:30 and 1:40; at 2-1/2 hours, at 2:15, 2:30 and 2:45; and at 3-1/2 hours, at 3:15, 3:30 and 3:45. Thus, each line consists of counts at independent and nonoverlapping time intervals. This method reduces variation, since the use of more time intervals before and after the mean time increases the variation around the mean time. The equations for the regression lines calculated are as follows:

| Mean time | Equation |
|-----------|--------------------|
| 1.5 hours | $y = 67.7 + 27x$ |
| 2.5 " | $y = 72.9 + 26.8x$ |
| 3.5 " | $y = 77.3 + 23.1x$ |

where y is the per cent of fish reacting and x is the coded log dose.

² Goldfish are now the subjects of a similar study; preliminary data follow. Ten goldfish one to two inches in length are placed in one-liter graduate cylinders containing one liter of distilled water each. A slow steady stream of air is bubbled through each cylinder. The water in the cylinders is brought to the same temperature as the water in the tanks from which the fish are taken. After allowing the fish to remain in the cylinders from 15 minutes to one-half hour, experiments are begun. The experiments are run for one hour and 10 minutes.

Concentrations of 0.1, 0.2, 0.5 and 1.0 mcg LSD-25 per ml have been tested on the goldfish. In concentrations of 0.5 and 1.0 mcg, in 10 minutes all of the fish are swimming at the top of the cylinder, nose up—tail down, though there is some swimming away from the top and returning. This pattern persists for the duration of the experiment. In the lower concentrations, there is more swimming away from the top, and the fish seem to "hang" there less. In concentrations of 0.2 mcg LSD-25 per ml, it takes 20-30 minutes to reach a maximum reaction of 70-100 per cent of the fish. More experiments need to be done on 0.1 mcg LSD-25 per ml. Controls occasionally have 20-30 per cent of the fish swimming at the top, but this has been reduced by increasing the waiting time at the beginning of the experiment.

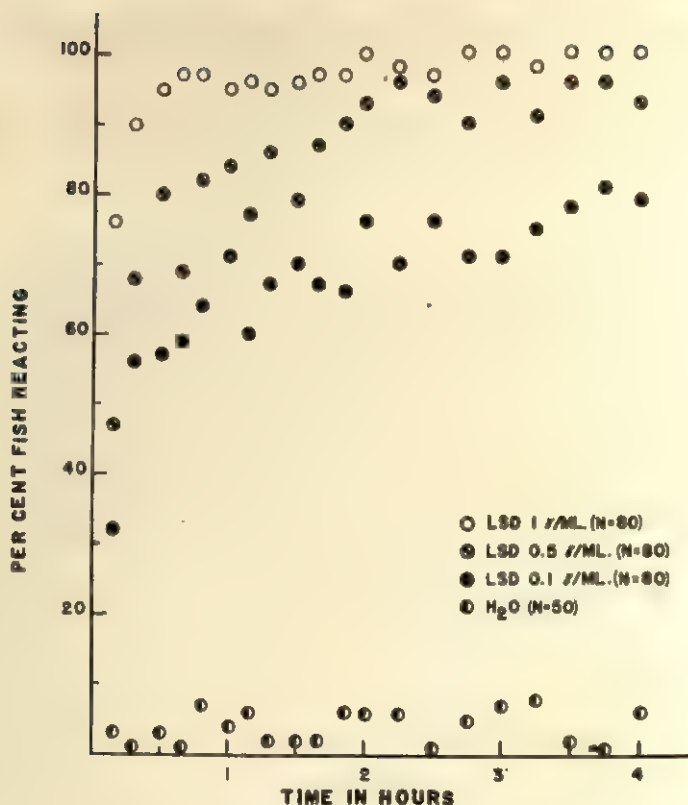


FIGURE 1

THE PER CENT OF *B. SPLENDENS* REACTING TO VARIOUS CONCENTRATIONS OF LSD-25 AS A FUNCTION OF TIME

The fish were young adults, placed in 100 ml of LSD-25 solution in the outside liquid. Each point represents an average of 80 fish at a given time, except in the case of controls where only 50 fish were used. The first reading was taken after 10 minutes. Temperature: 78° to 80° F.

Although the slopes at times 1-1/2 hours and 2-1/2 hours are quite parallel, the equation of the reaction at 3-1/2 hours probably represents the maximal effect of LSD-25 under our conditions. The per cent reacting under these conditions should be more stabilized than at the earlier counts. Calculation of the variance confirms this view.

When the fish are injected into the body cavity with LSD-25, the same general phenomena are observed. The fish react in the same way as they do when LSD-25 diffuses into the fish, presumably through the gills. However,

the reaction-time curve is different under our experimental conditions. The fish react to a much smaller amount of LSD-25 when injected. Thus, 0.05 ml of liquid containing only 0.01 mcg of LSD-25 can readily be detected. The per cent of fish reacting to injected LSD-25 with time is shown in Figure 2. The points are averages of 40 or more fish. When the fish are

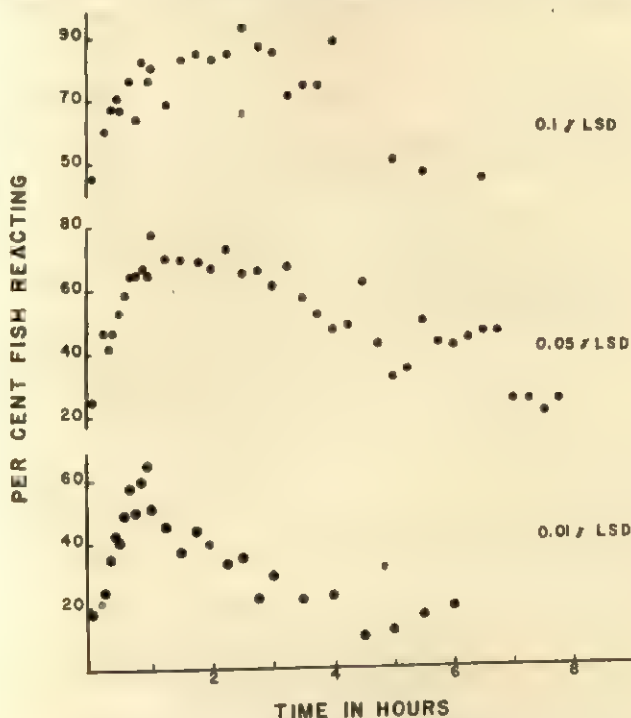


FIGURE 2

THE PER CENT OF *B. SPLENDENS* REACTING TO VARIOUS CONCENTRATIONS OF INJECTED LSD-25 AS A FUNCTION OF TIME.

LSD-25 was injected into the body cavity in a volume of 0.05 ml. The concentrations of LSD-25 injected were 0.01 ($N=70$), 0.05 ($N=50$) and 0.1 ($N=40$). A control group of 40 fish were injected with 0.05 ml H_2O . Each point represents the mean per cent reacting at that time. Zero time was counted as the start of injecting the first fish; 10 minutes is 10 minutes after the zero time.

injected with LSD-25, there is no longer a large reservoir of material in the outside liquid. Judging by the rapidity with which the LSD-25 acts when injected, equilibrium is more rapidly achieved. As the averaged raw data in Figure 2 illustrates, a maximum is obtained with the per cent of fish showing the typical LSD-25 reaction, which then decreases with time. By

studying the maximum response shown for various concentrations injected and obtaining a regression equation, we have found that the regression line for the maxima has a slope, within the limits of error, identical with the slopes observed at two hours when the LSD-25 is in the outside liquid. Thus, the regression equation for a mean time of two hours, with the LSD-25 in the outside liquid, is $y = 73 + 25.6x$, while the regression equation calculated

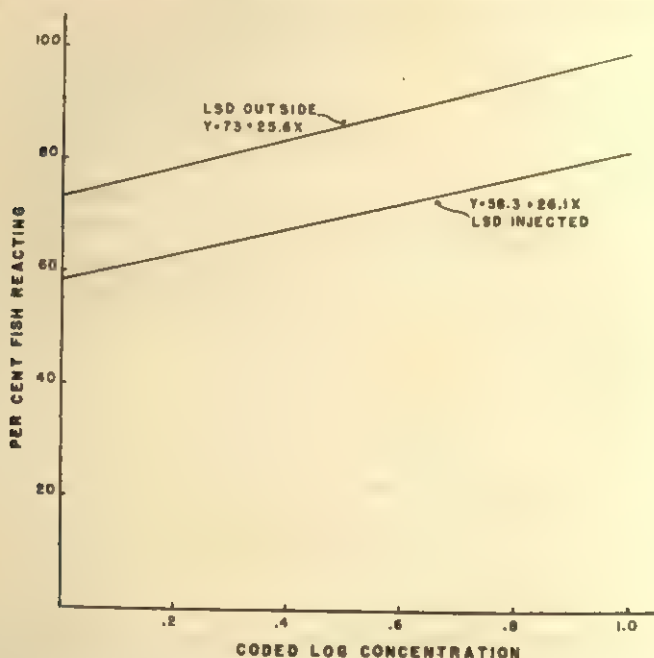


FIGURE 3

COMPARISON OF THE EFFECT OF LSD-25 ON THE SIAMESE FIGHTING FISH IN THE OUTSIDE LIQUID AND INJECTED DIRECTLY INTO THE BODY CAVITY OF THE FISH

The slopes of the regression lines are almost identical. Since the threshold is much lower for the injected LSD-25, the coding of the log of the concentration has enabled both to be plotted with the origins on the same ordinate. (Code for injected: $\log C + 2$; for LSD-25 outside: $\log C + 1$.) The time is two hours in the outside liquid and approximately the same mean time when injected.

for the injected fish at approximately the same mean time is $y = 58.3 + 26.1x$ (See Figure 3). There is a slight error in assuming that the maxima are at the time of two hours. Actually, it will be shown that there is a slight difference between the slopes of the regression lines for LSD-25 diffusing and injected. However, there will be closer parallelism between injected LSD-25 and injected Psilocybin. It would appear that LSD-25 acts similarly

whether it enters the gills or is injected into the body cavity. Important observations on the effect of LSD-25 on carp and game fish are reported by Loeb (5). It will appear from the data on Psilocybin in the section following that compounds simpler in structure than LSD-25 may also cause surfacing behavior in fish. It is most important, therefore, to extend the area of investigation to compounds similar in structure to Psilocybin when surfacing behavior is studied from the point of view of game-fish conservation.

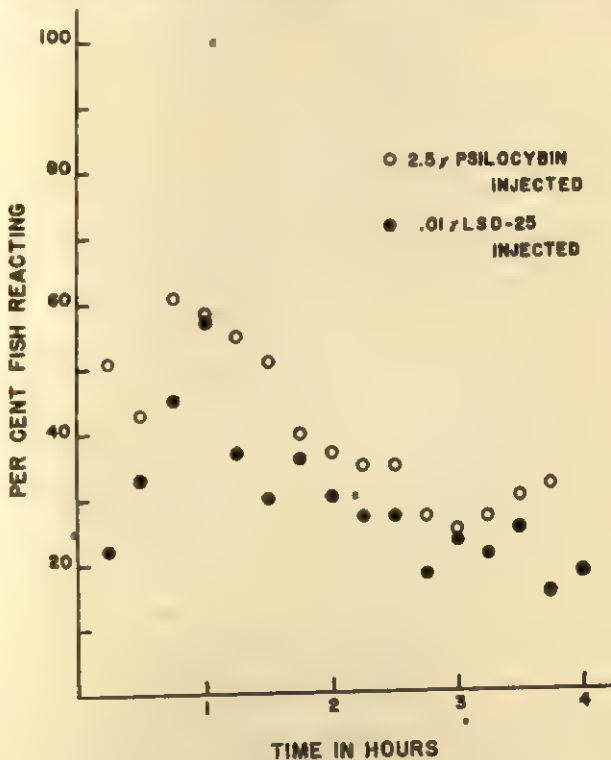


FIGURE 4
COMPARISON OF PER CENT OF FISH REACTING WHEN INJECTED WITH LSD-25 AND Psilocybin AT THE CONCENTRATIONS NOTED IN THE FIGURE

2. Part II: Psilocybin

Using the identical methods employed for the injection of LSD-25, fish were injected with Psilocybin in order to compare the reaction to Psilocybin to LSD-25. The nose up—tail down position and the other phenomena observed with LSD-25 are also observed in the fish injected with Psilocybin.

In addition, at very low doses (1.25 mcg) of Psilocybin the fish are apt to be flat at the surface, the angular position being reached only at higher concentrations.

The quantities of Psilocybin injected were 2.5 mcg ($N = 80$), 12.5 mcg ($N = 60$) and 25 mcg ($N = 60$). Six separate control experiments were run

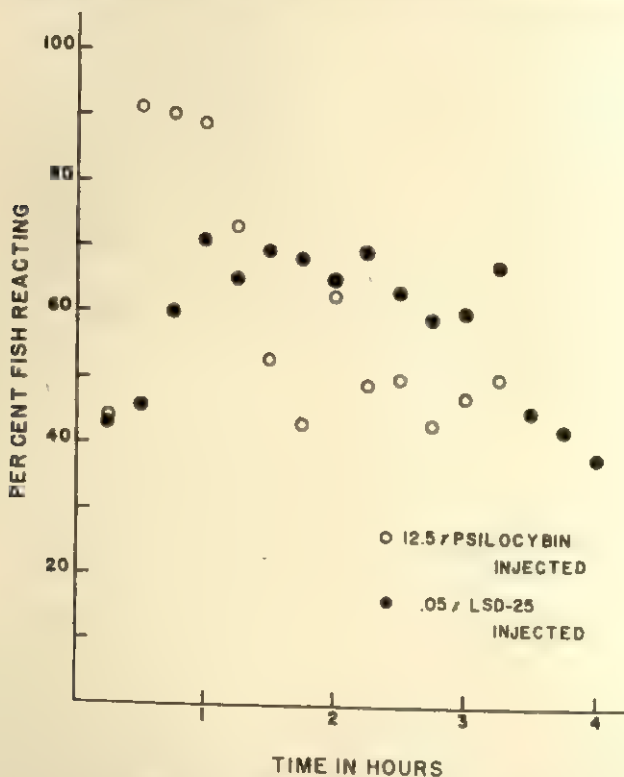


FIGURE 5
COMPARISON OF PER CENT OF FISH WHEN INJECTED WITH LSD-25 AND PSILOCYBIN AT THE CONCENTRATIONS NOTED IN THE FIGURE

with $N = 60$ fish in water controls. Although our data start with 2.5 mcg of Psilocybin injected, experiments with 1.25 mcg injected also show a reaction with most of the fish in a flat position at the surface. Unfortunately, when Psilocybin is placed in the outside liquid in concentrations as high as 100 mcg per ml, there is no reaction similar to LSD-25; thus, all of our data deal with injected Psilocybin.

Note in Figures 4, 5 and 6 a comparison of the per cent of fish reacting (nose up—tail down and Cartesian diver) with pharmacologically similar concentrations of LSD-25 and Psilocybin. It is evident that Psilocybin and LSD-25 are remarkably similar in their effects with, however, LSD-25 apparently producing a more rapid reaction but with the effect also decreasing

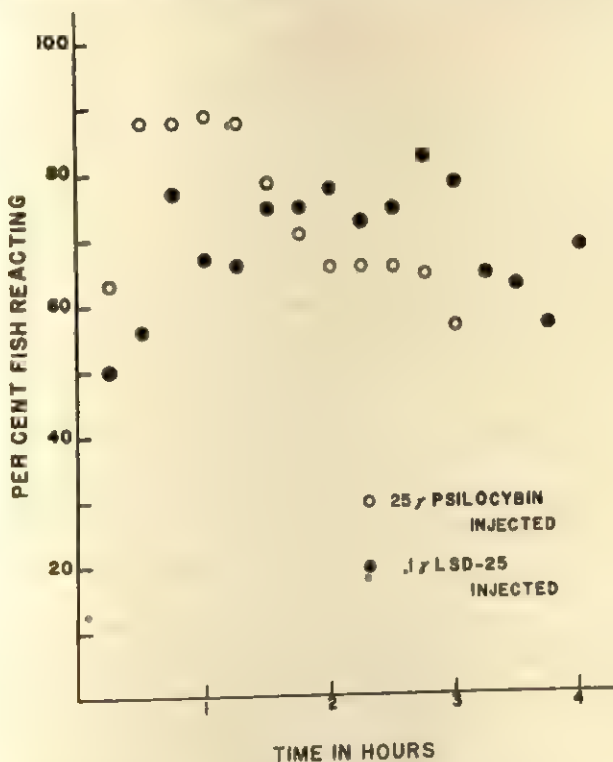


FIGURE 6

COMPARISON OF PER CENT OF FISH REACTING WHEN INJECTED WITH LSD-25 AND PSILOCYBIN AT THE CONCENTRATIONS NOTED IN THE FIGURE

more rapidly. This would, of course, have to do with the diffusion and metabolism of these drugs, which are different in chemical constitution. Figure 7 shows the regression lines comparing LSD-25 injected with Psilocybin injected. The slopes of 18.4 for LSD-25 and 19.4 for Psilocybin are very close. We have chosen to put on this graph, for comparison, the slope of 23.1 for LSD-25 in the outside liquid, although this differs for LSD-25 outside in Figure 3 because the slope of 23.1 here represents the slope at 3-1/2 hours rather than 2 hours as indicated in Figure 3.

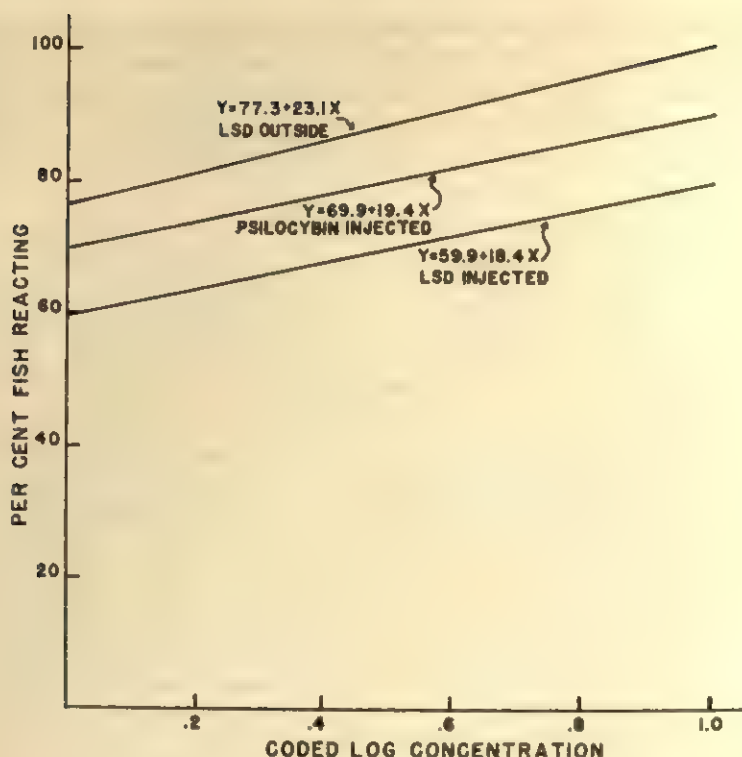


FIGURE 7

THE REGRESSION LINES COMPARING LSD-25 INJECTED WITH PSILOCYBIN INJECTED

The slopes of 18.4 for LSD-25 and 19.4 for Psilocybin are very close. We have chosen to put on this graph, for comparison, the slope of 23.1 for LSD-25 in the outside liquid although this differs for LSD-25 outside in Figure 3 because the slope of 23.1 here represents the slope at 3-1/2 hours rather than 2 hours as indicated in Figure 3.

D. DISCUSSION

Although systematic studies of the effects of psychotomimetic compounds on man and animals are being conducted in different laboratories, the use of the psychotomimetic agents on animals is fraught with considerable difficulty. Special responses of the central nervous system are readily studied, and isolated tissues are also investigated with facility. However, there are few small animals whose behavioral, vegetative and motor reactions can be simultaneously studied in large numbers. Our data indicate that the Siamese fighting fish is one type of animal which is suitable and readily available in standard populations for studies of this type. The argument may be made

that the fish is certainly far from man in the phylogenetic scale. However, our preliminary data indicate that the fish is almost as sensitive to LSD-25 as man is. Thus, a marked reaction to injected LSD-25 is obtained with only 0.01 mcg of LSD-25 in fish weighing one gram. The fish react, therefore, to 10 mcg pre kg of fish. This compares with a similar reaction of approximately 1.0 mcg per kg in man. Considering that the fish will react to less than 0.01 mcg of LSD-25, and that it is a cold-blooded animal, the sensitivity of the Siamese fighting fish is not very far from man. There are other advantages in using the fish as a test animal. Many reactions are slowed down in cold-blooded animals so that the intermediate steps of phenomena not readily observed in warm-blooded animals will manifest themselves in cold-blooded animals. In addition, there is justification for exploring enzyme mechanisms in lower forms like the Siamese fighting fish. It is well known that during the evolutionary process specific enzymes themselves do not often undergo extraordinary changes. Since both LSD-25 and Psilocybin effects can be mimicked in the fish by oxidase poisons like sodium cyanide and sodium azide (6), experiments in warm-blooded animals are directly suggested. It is believed that the integration of data obtained in fish with that obtained in man will lead to some understanding of the mechanisms of action of certain psychotomimetic drugs involved in man and higher animals. This path it is hoped may lead to understanding of clinical schizophrenia (4).

E. SUMMARY

The reactions of Siamese fighting fish to LSD-25 and Psilocybin disclose that injected Psilocybin affects the Siamese fighting fish in the same way that LSD-25 does. The slopes of the regression lines for LSD-25 injected and Psilocybin injected are almost identical. In all likelihood Psilocybin and LSD-25 act on the same enzyme systems. This is in keeping with observations in man that the reaction of man to Psilocybin is indistinguishable from that of LSD-25.

REFERENCES

1. ABRAMSON, H. A. Lysergic Acid Diethylamide (LSD-25): XXIX. The response index as a measure of threshold activity of psychotropic drugs in man. *J. of Psychol.*, 1959, **48**, 65-78.
2. ABRAMSON, H. A., & EVANS, L. T. Lysergic Acid Diethylamide (LSD-25): II. Psychobiological effects on the Siamese fighting fish. *Science*, 1954, **120**, 990-991.
3. ABRAMSON, H. A., ROLO, A., SKLAROFKY, B., & STACHE, J. Production of cross-tolerance to psychosis-producing doses of Lysergic Acid Diethylamide and Psilocybin. *J. of Psychol.*, 1960, **49**, 151-154.
4. ABRAMSON, H. A., SKLAROFKY, B., BARON, M. O., & GETTNER, H. H. Lysergic Acid Diethylamide (LSD-25) antagonists: I. Blocking effect of brain extract

- in the Siamese fighting fish; preliminary report. *Arch. Neurol. & Psychiat.*, 1960, **77**, 439-454.
5. LOEB, H. A. Effect of Lysergic Acid Diethylamide (LSD-25) on the surfacing behavior of laboratory carp. *N. Y. Fish & Game J.*, 1962, **9**, 127-132.
 6. WEISS, B., ABRAMSON, H. A., & BARON, M. O. Lysergic Acid Diethylamide (LSD-25): XXVII. Effect of potassium cyanide and other oxidase and respiratory inhibitors on the Siamese fighting fish. *Arch. Neurol. & Psychiat.*, 1958, **80**, 345-350.

South Oaks Research Foundation, Inc.
Sunrise Highway
Amityville, Long Island, New York

EMOTIONAL BEHAVIOR OF THE WHITE RAT FOLLOWING FULL-THICKNESS THERMAL BURNS*¹

Departments of Surgery and Pediatrics, Upstate Medical Center, Syracuse, New York

FRED T. CALDWELL, JR., BETTYE M. CALDWELL, AND STANLEY
E. MOLDOVAN

A. INTRODUCTION

The physician called upon to treat burns is frequently handicapped by symptoms of emotional disturbance in the patient. In the medical literature one finds reports of such reactions as marked irritability, aggressiveness, overdependency, paranoid tendencies, and an assortment of devious behaviors which often succeed in subverting therapy.

Well-controlled clinical studies concerned with this problem are rare. Adler (1) and Cobb and Lindemann (5) have reported on two series of patients treated in two Boston hospitals following the disastrous Coconut Grove Fire in 1942. On the basis of a psychiatric examination, Cobb and Lindemann (5) concluded that 14 of 32 patients who survived presented neuropsychiatric problems. In addition to examinations during the immediate postburn period, Adler (1) had follow-up data on 46 cases. Of this group, 20 had shown no psychiatric complications at any time, 25 had had symptoms of anxiety and nervousness for approximately three months, while 13 still suffered from marked anxiety nine months after the disaster. Incidence of disturbance seemed unrelated to severity of burn.

In these two samples there is no need to consider the question of whether the emotional disturbance preceded the burn and may even have been a factor in the production of the trauma. However, Cobb and Lindemann (5) felt that all of the cases showing clear neurotic or psychotic reactions had histories suggesting the likelihood of such reactions under any severe stress. In a series of 19 burned children, Long and Cope (7) found evidence of significant maladjustment prior to the burn in half of the cases.

* Received in the Editorial Office on June 7, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ This study has been supported in part by Grant RG-7151, U.S. Public Health Service.

² The assistance of Mrs. Elaine Erichson and Mr. Howard Crane in various aspects of this project is gratefully acknowledged.

This question of whether burns produce emotional disturbance or emotional problems tend to produce burns has been investigated by Woodward (8) in a study of 198 children under 15 years of age who sustained a burn of 10 per cent or more of body-surface area. Two to five years after the accident, she found that 81 per cent of the children showed some degree of emotional disturbance, with 67 per cent rated as moderately or severely disturbed. In two control groups consisting of 608 siblings of the burned children and 50 randomly selected children from comparable sociogeographic districts, incidence of disturbance was 7 per cent and 14 per cent, respectively. Fears and anxieties were the outstanding problems reported for the burned children, followed in order by feeding difficulties, sleep disorders, and enuresis. The author concluded that there was little evidence to suggest that the children were emotionally disturbed prior to the trauma, although there was a slight tendency for children who showed signs of preburn disturbance to become severely disturbed after the accident.

Long and Cope (7) suggest that emotional disturbance following a burn wound might be accounted for by the psychological impact of the trauma, the hormonal and metabolic response of the organism, or postburn infection. As they found severe emotional reactions in cases with small burn, no fever, and minimal metabolic reaction, they concluded that threat implied in the trauma and enforced hospitalization outweighed the other factors. Adler's finding (1) that those patients who lost consciousness were slightly more likely to be symptom-free than those who remained conscious might be interpreted as offering evidence on this question.

The metabolic and nutritional readjustment required of the burned organism is profound; likewise, patterns of securing contact comfort are often drastically altered. Such alterations in the interaction between the animal and its physical and social environment may be of such magnitude as to foster aberrant behavior. In a series of investigations (2, 3, 4) concerned with physiologic response to thermal trauma in the white rat, one of the authors (FTC) had observed occasional dramatic postburn changes in the behavior of the experimental animals. As the animals were anesthetized for the production of the burn wounds, the impact of the trauma itself did not seem adequate to account for the observed changes. Nor could the possible cosmetic injury or psychological meaning of the event be implicated. Accordingly the present study was set up as a controlled test of the experimenter's "clinical impressions." By an investigation of this problem in the white rat, evidence would be made available on the question of whether there is something inherent in the burn wound, or possibly any injury involving significant

loss of skin covering, which predisposes to a change in emotional functioning.

The design of the study permitted phrasing the null hypothesis as follows: There will be no difference in emotional behavior, operationally defined, between a group of animals with full-thickness burns covering approximately 20 per cent of body surface and a comparable group of unburned controls.

B. PROCEDURE

1. Subjects

The experimental sample consisted of 48 male Sprague-Dawley strain rats divided randomly upon arrival in the laboratory into experimental and control animals. Twenty-eight animals were assigned to the experimental (burn) group and 20 to the control group in anticipation of some mortality in the burn group. Four experimental animals expired on the day of the burns and one additional animal died 45 days postburn; one control animal died 41 days postburn. Thus the two groups at the termination of the experiment had 23 and 19 animals. All animals were kept in individual cages stacked on a large portable rack which was systematically rotated to control for any minor light and temperature variations within the laboratory. Ambient temperature was maintained at 20°C. Food and water were supplied *ad libitum*, and daily intake of each was recorded throughout the experimental period. All animals were weighed once a week.

2. Experimental Burns

Experimental burns on the backs and flanks of the animals were produced by immersion for 30 seconds in water of 90°C. The resultant burn wounds had a mean size of 21 ± 2 per cent of the body surface area. All animals, both burn and nonburn, received between 100 and 120 seconds of ether anesthesia.

3. Behavioral Measures

a. Activity test. The apparatus for measuring nocturnal activity consisted of a standard activity wheel. Activity was measured for a 14-hour period under conditions of total darkness for eight hours. As six activity cages were available, each animal was run approximately once a week for six weeks. A table of random numbers was used to establish the sequence in which the animals were run. Because of time and laboratory space limitations, measures of activity had to be discontinued after six weeks, although all other measures were made for 13 weeks.

b. Open field test. This is probably the most frequently used measure of

emotionality in the rat. In standardizing their procedure, the experimenters adapted the procedure described by Hall (6). The open field constructed for this investigation consisted of a white plywood box 35 inches square with walls 13 inches high. The floor area was divided into smaller squares 7" \times 7", thus making 16 peripheral and nine inner squares. The box was placed on a rotating laboratory stool and systematically turned between runs. The floor of the box was covered with linoleum and washed between runs in order to minimize olfactory tracking cues. During the 90-day experimental period, three degrees of stress in the field were arranged. For the first 15 runs, moderate stress was arranged by placing the box directly under a bright light and attaching to the shaft of the supporting stool a loud doorbell buzzer. Then during Trials 16-18 the stress was reduced by eliminating the buzzer. In Trials 19 and 20 the stress was greatly intensified by giving each animal three five-second periods of shock (90 volts) during each run, making a total of 30 seconds of shock. For these increased stress trials, the floor of the box was covered with an electrified grill. On the final two trials, the original bell-light conditions were reintroduced. The open field was varied in this manner in order to determine whether burned and non-burned animals would respond differently to variations in level of external stress.

Emotionality in the open field was defined as lack of exploratory behavior, preference for outer and avoidance of inner squares, relatively high incidence of defecation and urination in the box, and inability to adapt to the apparatus after repeated exposure.

Data were analyzed in terms of between-group comparisons using either the *t* test or chi square. The .05 level of probability was accepted in advance as the level which would be required for rejecting the null hypothesis.

C. RESULTS

Results of the experiment are presented in five time blocks as follows:

- I. Preburn period (one week preburn).
- II. Early postburn period, from the day of burn to the time of separation of the eschar (1st to 31st postburn day).
- III. Period of reduced environmental stress (32nd to 40th postburn day).
- IV. Period of increased environmental stress (48th to 50th postburn day).
- V. Reprise period, after no exposure of animals to experimental handler or open field for 30 days (80th to 87th postburn day).

In Table 1 are presented data pertaining to food and water intake and

TABLE 1
FOOD AND WATER INTAKE AND WEIGHT MEASURES FOR BURNED AND CONTROL ANIMALS

| Time period | Burn group | | | Control group | | | Diff. | t | p |
|-------------------------------|------------|-------|------|---------------|-------|------|-------|-------|-------|
| | N | M | SD | N | M | SD | | | |
| Food intake (in grams) | | | | | | | | | |
| I | 28 | 27.2 | 2.2 | 20 | 27.5 | 1.5 | .3 | .57 | — |
| II | 24 | 31.7 | 1.8 | 20 | 31.6 | 1.7 | .1 | .60 | — |
| III | 24 | 32.4 | 3.2 | 20 | 30.0 | 2.0 | 2.4 | 3.03 | .005 |
| IV | 23 | 34.6 | 3.6 | 19 | 31.9 | 2.0 | 2.7 | 3.10 | .005 |
| V | 23 | 34.5 | 3.3 | 19 | 29.9 | 1.8 | 4.6 | 5.22 | <.001 |
| Water intake (in milliliters) | | | | | | | | | |
| I | 28 | 45.3 | 5.6 | 20 | 46.4 | 4.3 | .9 | .63 | — |
| II | 24 | 65.8 | 5.8 | 20 | 48.9 | 4.5 | 16.9 | 10.9 | <.001 |
| III | 24 | 70.8 | 9.6 | 20 | 45.4 | 6.0 | 25.4 | 10.7 | <.001 |
| IV | 23 | 73.9 | 11.1 | 19 | 47.1 | 4.9 | 26.8 | 10.47 | <.001 |
| V | 23 | 73.9 | 12.6 | 19 | 45.4 | 5.0 | 28.5 | 10.0 | <.001 |
| Weight (in grams) | | | | | | | | | |
| I | 28 | 254.1 | 10.9 | 20 | 253.6 | 8.5 | .5 | .02 | — |
| II | 24 | 332.1 | 17.4 | 20 | 352.8 | 15.4 | 20.7 | 4.04 | <.001 |
| III | 24 | 359.0 | 26.4 | 19 | 400.1 | 20.7 | 41.1 | 5.56 | <.001 |
| IV | 23 | 337.6 | 22.9 | 19 | 425.6 | 21.9 | 48.0 | 6.89 | <.001 |

weight gain during the entire experimental period. From the 32nd postburn day to the end of the experiment the burn group consumed significantly more food than the control group. This increment in daily food consumption is related to an obligatory energy loss from the burn wound resulting in evaporative heat loss.

The increment in vaporizational water loss is reflected in the sustained and significant increment in water consumption by the burn group which started immediately after burning and continued throughout the experiment.

The body-weight data demonstrate that growing albino rats do not resume a normal growth rate for at least 80 days following thermal injury, even when allowed free access to food at all times. The mean body weight of the burned group was significantly below that of the nonburned group for all postburn measurements. These data confirm well-established nutritional and metabolic responses of the albino rat to thermal injury at an ambient temperature of 20°C (2, 3, 4).

Table 2 presents the results of activity cage performance of the two groups of animals. Unfortunately activity cage data are available for only three of the time periods. There is a significant difference between the activity level of the burned and control animals during Period II, the immediate postburn period. This difference can undoubtedly be attributed to motility limitations imposed by the burn wound during that period and by

TABLE 2
NOCTURNAL ACTIVITY CAGE PERFORMANCE OF BURNED AND CONTROL ANIMALS

| Time period | Burn group | | | Control groups | | | Diff. | <i>t</i> | <i>p</i> |
|-------------|------------|----------|-----------|----------------|----------|-----------|-------|----------|----------|
| | <i>N</i> | <i>M</i> | <i>SD</i> | <i>N</i> | <i>M</i> | <i>SD</i> | | | |
| I | 19 | 427.8 | 178 | 16 | 444.7 | 164 | 16.9 | .28 | — |
| II | 24 | 427.3 | 172 | 20 | 685.5 | 226 | 258.2 | 4.20 | .001 |
| III | 19 | 342.9 | 337 | 15 | 480.9 | 132 | 138 | 1.45 | — |

the suboptimal nutritional state of the animals. The control animals continued to show more mean activity during Period III, but the difference between the groups did not attain statistical significance.

Table 3, 4, and 5 present the data from the open-field test. Hall (6) offers a persuasive argument that the tendency to urinate or defecate while in the open field is indeed a sign of emotionality in the rat. In the present experiment, there were no differences between burned and nonburned

TABLE 3
EMOTIONALITY INDICES DURING OPEN-FIELD TEST FOR BURNED AND CONTROL ANIMALS

| | Urinations | | | | Fecal boli | | | |
|---------|------------|-------|-------------|----------|------------|-------|-------------|----------|
| | 1 or more | None | \bar{X}^a | <i>p</i> | 1 or more | None | \bar{X}^a | <i>p</i> |
| I | | | | | | | | |
| Burn | 8 | 20 | 1.44 | — | 16 | 12 | .29 | — |
| Control | 2 | 18 | | | 9 | 11 | | |
| II | | | | | | | | |
| Burn | 11 | 13 | .18 | — | 7 | 17 | .07 | — |
| Control | 7 | 13 | | | 6 | 14 | | |
| III | | | | | | | | |
| Burn | 3 | 21 | .05 | — | 5 | 19 | .32 | — |
| Control | 2 | 18 | | | 2 | 18 | | |
| IV | | | | | | | | |
| Burn | > Mdn | < Mdn | .16 | — | > Mdn | < Mdn | .38 | — |
| | 11 | 12 | | | 13 | 10 | | |
| Control | 7 | 12 | | | 8 | 11 | | |
| V | | | | | | | | |
| Burn | 11 | 12 | .47 | — | 9 | 14 | .01 | — |
| Control | 12 | 7 | | | 8 | 11 | | |

animals in either of these indices of emotionality at any of the time periods. Repeated exposure of the animals to the field should reduce the likelihood of either urination or defecation, and the decreasing tendency of animals from either group to manifest these behaviors during Periods I, II, and III indicates

TABLE 4
MOTILITY SCORES OF BURNED AND CONTROL ANIMALS DURING OPEN-FIELD TEST

| Time period | Burn group | | | Control group | | | Diff. | t | p |
|-----------------------------------|------------|------|------|---------------|------|------|-------|------|------|
| | N | M | SD | N | M | SD | | | |
| Number of outer squares traversed | | | | | | | | | |
| I | 28 | 69.6 | 18.4 | 20 | 68.6 | 14.0 | 1.0 | .19 | — |
| II | 24 | 39.8 | 12.8 | 20 | 43.7 | 19.5 | 3.9 | .77 | — |
| III | 24 | 30.3 | 16.1 | 20 | 31.7 | 17.6 | 1.4 | .26 | — |
| IV | 23 | 33.8 | 8.5 | 19 | 33.8 | 8.8 | 0.0 | — | — |
| V | 23 | 21.2 | 12.9 | 19 | 8.4 | 5.6 | 12.8 | 3.94 | .001 |
| Number of inner squares traversed | | | | | | | | | |
| I | 28 | 5.5 | 4.1 | 20 | 3.9 | 1.8 | 1.6 | 1.62 | — |
| II | 24 | 5.3 | 2.8 | 20 | 5.11 | 2.80 | .2 | .09 | — |
| III | 24 | 6.9 | 4.3 | 20 | 4.6 | 2.4 | 2.3 | 2.14 | .05 |
| IV | 23 | 5.4 | 2.3 | 19 | 4.6 | 2.3 | .8 | 1.01 | — |
| V | 23 | 4.2 | 2.5 | 19 | 3.5 | 1.6 | .7 | 1.09 | — |
| Number of rises | | | | | | | | | |
| I | 28 | 21.2 | 8.9 | 20 | 21.0 | 5.5 | .2 | .10 | — |
| II | 24 | 11.3 | 5.9 | 20 | 14.7 | 5.2 | 3.4 | 1.97 | — |
| III | 24 | 7.7 | 4.2 | 20 | 9.4 | 4.8 | 1.7 | 1.24 | — |
| IV | 23 | 3.6 | 2.2 | 19 | 3.5 | 1.7 | .1 | .21 | — |
| V | 23 | 7.2 | 4.3 | 19 | 3.6 | 2.7 | 3.6 | 3.05 | .01 |
| Squares covered during shock | | | | | | | | | |
| IV | 23 | 29.3 | 7.3 | 19 | 23.6 | 10.1 | 5.7 | 2.07 | .05 |

TABLE 5
ESCAPE ATTEMPTS IN OPEN FIELD DURING AND FOLLOWING SHOCK

| | Shock period (IV) | | | | Postshock period (V) | | | |
|---------|-------------------|----------------|----------------|---|----------------------|----------|----------------|-----|
| | Jumps > Mdn. | Jumps < Mdn | X ² | p | Jumps | No jumps | X ² | p |
| Born | 10 | 13 | .92 | — | 2 | 21 | 4.56 | .05 |
| Control | 12 | 7 | | | 8 | 11 | | |

that this reduction occurred in both groups. During Period IV, the shock period, all animals urinated one or more times in the field, and most animals defecated. Therefore for Periods IV and V the median test was used for obtaining the chi squares. No significant differences in the behavior of the animals in the two groups could be observed, however.

In Table 4 the motility scores of the animals in the open field are presented. In terms of the operational criteria accepted in advance, results suggest that the burned animals were, if anything, less emotional than the unburned controls. During the low-stress period (Period III) the burned animals moved about more on the inner squares than the control animals. But the most striking differences were during Period V—the reprise period

one month after the shock trials. Most of the control animals, when put back into the box one month after the two shock trials, seemed almost as fearful of the box as they had been at the time of the second shock trial. The wire floor had been removed for these last trials so that the visual and tactile stimulation associated specifically with the shock were now gone and only the "safe" properties were then perceptible. Yet many of the control animals seemed to react with the full intensity shown during the shock trials. The burned animals, on the other hand, seemed unperturbed by being put back in the box. Although their general motility (outer plus inner squares) was somewhat reduced, they continued to move about in the inner squares and resumed their leisurely and investigatory rises in the box.

During the shock period a new response measure appeared—jumping out of the box to escape the shock. During the shock trials, animals in the two groups distributed themselves randomly about the median number of jumps (Table 5). However, during the postshock period, only two of the 23 burned animals jumped in the apparatus, while eight of the 19 controls did so. It was as though the burned animals either did not learn to fear the box, or else that they quickly shed their fear. An alternative interpretation might be that they were perhaps more alert and learned more efficiently, as, upon encountering the resumption of the preshock stimulus configuration in Period V, they resumed the more appropriate pattern of preshock behavior. It is also possible that the burned animals did not get quite as much shock as the control animals. The last line of Table 4 shows that the burned animals moved around during the shock more than the control animals did, though during the 165 seconds of each run when the shock was off, they were no more active than the controls. It is conceivable that their additional movements somehow reduced the intensity of the shock they received.

D. DISCUSSION

The results of this experiment did not reveal any striking increase in emotionality in the burned rat in contrast to comparable unburned controls. If anything, the burned animals became less emotional in comparison to the controls over time. This was especially striking in Period V, the postshock period, during which the control animals gave every indication of having learned to fear the open field and the experimental situation while the burned animals showed almost no residual fear. It is difficult to account for this difference in terms of any direct sequelae of the wound. If the burn had seriously hampered the motor ability of the burned animals, one would not expect them to remain on a par with the controls in the rudimentary type

of locomotion involved in open-field movement. During the shock trials the burned animals easily demonstrated their ability to execute the somewhat difficult jump out of the box; also during Period V they were entirely able to move about and to rise up on two feet. Rather it seems that there was a real difference in reactivity between the groups at the end of the experiment, with the burned animals showing reduced emotionality in the open field.

E. SUMMARY

In evaluating the clinical observation that recovery from burns in humans is frequently complicated by emotional reaction, it is difficult to separate the role of the social and psychological implications of the burn wound from the dramatic physiological adaptation necessitated by the thermal injury. Accordingly, an experimental investigation, using the white rat as the animal of choice, was conducted in which severity of the wound could be controlled, treatment managed without drastic alteration of the social environment of the wounded animal, and the etiologic force associated with personal disfigurement presumably eliminated. Twenty-four burned and 20 control animals were intermittently put through a six-week series of activity measures and a three-month series of open-field tests with systematic variation in level of stress. Behavioral differences between the groups were slight until the last test period, which followed a shock-test period by one month. At that time the burned animals showed significantly less freezing in the apparatus, more relaxed investigatory behavior, and made fewer attempts to escape the formerly painful apparatus. These differences were interpreted in terms of a decreased incidence of learning to fear the apparatus on the part of the burned animals in contrast to the nonburned controls. Although the meaning of this difference is not clear in terms of the adaptive behavior of the burned animals, the general findings of the study suggest that the psychosocial implications of the burn wound in humans may well outweigh the physiologic changes to which the body must adjust in influencing the behavioral response.

REFERENCES

1. ADLER, A. Neuropsychiatric complications in victims of Boston's Coconut Grove disaster. *J. A. M. A.*, 1943, **123**, 1098-1101.
2. CALDWELL, F. T., OSTERHOLM, J. L., SOWER, N. D., & MOYER, C. A. Metabolic response to thermal trauma of normal and thyropivic rats at three environmental temperatures. *Ann. Surg.*, 1959, **150**, 976-988.
3. CALDWELL, F. T. The effect of early eschar excision upon the nutritional state and rate of wound closure of rats with full-thickness thermal burns. *Surgery*, 1961, **49**, 454-460.
4. ———. Metabolic response to thermal trauma: II. Nutritional studies with rats at two environmental temperatures. *Ann. Surg.*, 1962, **155**, 119-126.

5. COBB, S., & LINDEMANN, E. Symposium on management of Coconut Grove burns at Massachusetts General Hospital: Neuropsychiatric observations. *Ann. Surg.*, 1943, **117**, 814-824.
6. HALL, C. S. The inheritance of emotionality. In W. E. Martin & C. B. Stendler (Eds.), *Readings in Child Development*. New York: Harcourt Brace, 1954. Pp. 59-68.
7. LONG, R. T., & COPE, O. Emotional problems of burned children. *New Eng. J. Med.*, 1961, **264**, 1121-1127.
8. WOODWARD, J. Emotional disturbances of burned children. *Brit. Med. J.*, 1959, **i**, 1009-1013.

College of Medicine
Department of Pediatrics
Upstate Medical Center
766 Irving Avenue
Syracuse 10, New York

TOWARD AN AUTOMATED TECHNIQUE OF PROBING INTO EMOTIONAL BLOCKS*

Department of Psychology, West Virginia Wesleyan College

ALBIN R. GILBERT

A. INTRODUCTION

In a preliminary research into social perception, this author found essentially that a normal, wholesome personality shows a consistent tendency to identify itself with "congenial" statements faster, and with "uncongenial" ones more slowly; and that exceptions to this rule disclose emotional blocks underlying the subject's reactions (7, 8).

The principle of causing a subject to "step into the shoes of" another person who is prompted by a strong motive, and then later "into the shoes of" another person actuated by the weak pole of the same motive is termed by this author the "principle of cross-examination." Studies on the differential reactions of subjects to *projective* "cross-examinative" stimulus propositions (4, 5, 6) were followed by an investigation of differential reaction times of a subject's accepting or rejecting "congenial" or "uncongenial" motivational setting, respectively (7, 8).

Research related to the use of reaction times as "verifications" of self-reports by subjects can be traced from C. G. Jung's classical Word Association Test (10, 11). Buytendijk analyzed the famous UFA films that show fighting scenes between men and animals (3). In a tennis match, for example, the opponents display faster reaction time in their interpersonal perception than they do in calm relationships.

While this observation reveals the *emergency-provoked shortening* of reaction time, the present research project is interested in the *anxiety-induced lengthening* of reaction time.

Some results obtained by the American "New Look" perception research are partly related to the technique proposed in the present paper (1, 2, 9, 12, 13). The bearing of this literature on the present problem was briefly discussed in a paper by this author, delivered in a symposium at the XVI International Congress of Psychology, Bonn, 1960 (7, 8).

* Received in the Editorial Office on June 11, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

B. PROCEDURE

Realizing that the intriguing phenomenon of differential reaction time in a subject's acceptance or rejection of "congenial" or "uncongenial" contrasting motivational settings of another person may possibly lead to a novel effective technique of personality assessment and, especially, to the diagnosis of specific anxieties, the author set up a technique for experimentation which works essentially as follows:

Let us suppose a self-reliant subject reads a statement appearing in a cameralike shutter, or hears it spoken by a tape recorder, such as: "I can go on with my work without encouragement by other people."¹ The subject indicates whether or not the proposition habitually applies to him by depressing one of two buttons, attached to a chronoscope. In this way, the subject's decision time from exposure to the card to the moment of decision is registered, concurrently with his "agree" or "disagree" signalled by the buttons. Sometime later, out of order, interrupted by unrelated items, a counter-item is brought into view: "I can't go on with my work without encouragement by other people." Notice the same reading time as in the item before. This counter-item reflects a lack of self-reliance. Again, the same measures are recorded.

The comparison of the reaction times to the "cross-items" reveals that a genuinely self-reliant subject, that is one whose self-reliance has no underlying emotional block, identifies himself with the "self-reliant item" faster, and with the "unself-reliant" one more slowly. It was found that exception to this normal reaction time pattern indicate areas of emotional disturbance, anxieties, "admixed to" the tested dimensions. The categories of these exceptions are explained later.

For the time being, the significance of this technique is being examined on a clinical basis by comparing and contrasting individual criterion persons, "strong" on mental health characteristics, with persons "weak" on the same characteristics. The weak group was drawn from patients of Brockton VA Hospital, diagnosed as neurotic or psychotic, or from maladjusted college students, identified as such by the college administration and faculty. The strong group was selected on the basis of the pooled judgment of persons well acquainted with the subjects.²

¹ Adapted from California Test of Personality—Adult, Form A, item 6, section 1A on Self-Reliance.

² The author takes this opportunity to thank Dr. Frederick H. Heilizer, research psychologist, VA Hospital, Brockton, and other members of the research staff, for making this research possible.

With a view to the subtle shadings of interpretations resulting from the differential reaction times in responding to "cross-examinative" stimulus items, this intensive clinical method of validation and evaluation is very revealing.

In a second phase of validation, a test of significance will be based on large criterion populations. The aim of this phase will be to obtain some comparative quantitative formulas, differentiating between criterion groups "high" on mental health traits as over against such as are "low" on the same traits.

To illustrate the clinical test of significance, let us consider the following abridged testing protocols of two criterion persons: subject A, a normal, mature person, against subject B, a neurotic patient.

These contrasting persons were subjected to "timed cross-examination" by the same set of diagnostic dimensions, patterned on mental health characteristics. As the purpose of the investigation was merely to prove the discriminatory value of "cross-examinative" stimulus sets, stimulus material from any reliable, validated, and standardized objective personality test can be drawn upon, provided that for each "borrowed" stimulus statement a counterstatement is prepared. Accordingly, a selection of test items from the very thoroughly developed, well-known *Mental Health Analysis*, published by the California Test Bureau, was adapted to the "cross-examinative pattern."

Table 1, for subject A and subject B, shows the same "dimensions by categories." Under each aspect of mental health, divided according to the above test into five aspects of Assets and five aspects of Liabilities, two sets of contrasting stimulus statements are listed. In the actual test instruments, more than two sets examine the subject on each dimensional category, as for example on the mental health aspects in the present illustration.

As we know, the two propositions of each set are so patterned as to test the subject's response to the strong and the weak pole, respectively. Thus, subject A's "close personal relationships," are tapped by sets 1, 2 and 3, 4 (Table 1, column 3). "Interpersonal skills" are tested by sets such as 5, 6 and 7, 8.

The list, "Dimensions by Categories," serves as a worksheet for the interpretation. The "Random List of Dimensions," however, not illustrated in this paper, contains the same stimulus items as the systematic sheet ("Dimensions by Categories," Table 1); but in the "Random List of Dimensions," the items are intermixed in random order. This list is used for the administration of the test.

TABLE 1
DIMENSIONS BY CATEGORIES

| (1) Numbers Random | (2) System- atic | (3) Dimensions | (4) + or — | (5) Speed of Agree | (6) Dis- agree | (7) + or — | (8) Inter- pretive patterns |
|--------------------------------------|------------------------|---|---------------------|--------------------------|----------------------|---------------------|--------------------------------------|
| <i>Subject A (wholesome, mature)</i> | | | | | | | |
| <i>Close personal relationships</i> | | | | | | | |
| 36 | 1 | If I am in trouble, there is someone to whom I can go. | + | 1.8 | | + | } Regular |
| 2 | 2 | There is no one to whom I can go if I am in trouble. | — | | 4.0 | + | |
| 27 | 3 | I have a number of friends among members of the opposite sex. | + | 2.6 | | + | } Regular |
| 14 | 4 | I have no friends among members of the opposite sex. | — | | 2.8 | + | |
| <i>Interpersonal skills</i> | | | | | | | |
| 20 | 5 | I usually show an interest in the things my acquaintances are doing. | + | 0.8 | | + | } Regular |
| 19 | 6 | I rarely show an interest in the things my acquaintances are doing. | — | | 3.4 | + | |
| 28 | 7 | It is easy for me to be friendly with people even when they disagree with me. | + | | 2.8 | — | } Reversed polarity |
| 8 | 8 | It is difficult for me to be friendly with people when they disagree with me. | — | 2.1 | | — | |
| <i>Social participation</i> | | | | | | | |
| 11 | 9 | I am a member of a club or other group where people do interesting things together. | + | | 2.0 | — | } Nonpolarity |
| 9 | 10 | I am not a member of a club or other group where people do interesting things together. | — | | 3.9 | + | |
| Bal. forward | | | | 7.3 | 18.9 | | |

TABLE 1 (continued)

| (1) Random | (2) Numbers System- atic | (3) Dimensions | (4) + or — | (5) Speed of Agree | (6) of Dis- agree | (7) + or — | (8) Inter- pre- tive patterns |
|---------------------------------------|-----------------------------------|--|---------------------|--------------------------|----------------------------|---------------------|---|
| <i>Subject A (wholesome, mature)</i> | | | | | | | |
| | | Bal. forward | | 7.3 | 18.9 | | |
| 22 | 11 | When I have time I call on my friends. | + | 2.6 | | + | } Regular |
| 29 | 12 | When I have time I don't call on my friends. | — | | 4.8 | + | |
| <i>Satisfying work and recreation</i> | | | | | | | |
| 35 | 13 | I usually look forward with pleasure to the duties of each new day. | + | | 2.6 | — | } Reversed polarity |
| 6 | 14 | I rarely look forward with pleasure to the duties of each day. | — | 3.5 | | — | |
| 38 | 15 | I have as much time for play and recreation as I should. | + | 2.3 | | + | } Regular |
| 37 | 16 | I have less time for play and recreation as I should. | — | | 2.9 | + | |
| <i>Adequate outlook and goals</i> | | | | | | | |
| 26 | 17 | I make plans for carrying on a successful career. | + | | 1.4 | — | } Reversed polarity |
| 3 | 18 | I don't make plans for carrying on a successful career. | — | 2.2 | | — | |
| 12 | 19 | I feel that there is some great plan in the universe and that I have a part in it. | + | 2.1 | | + | } Regular |
| 25 | 20 | I don't feel that there is some great plan in the universe and that I have a part in it. | — | | 2.5 | + | |
| <i>Behavioral immaturity</i> | | | | | | | |
| 13 | 21 | I find that I get along best when I concern myself about other people's feelings. | + | | 4.4 | — | } Nonpolarity |
| 10 | 22 | I find I get along best when I don't concern myself about other people's feelings. | — | | 2.4 | + | |
| | | Bal. forward | | 20.0 | 39.9 | | |

TABLE 1 (continued)

| (1) Random | (2) Numbers System- atic | (3) Dimensions | (4) + or — | (5) Speed of Agree | (6) Speed of Dis- agree | (7) + or — | (8) Inter- pretive patterns |
|-------------------------------|-----------------------------------|--|---------------------|--------------------------|----------------------------------|---------------------|--------------------------------------|
| Subject A (wholesome, mature) | | | | | | | |
| | | Bal. forward | | 20.0 | 39.9 | | |
| 31 | 23 | I have found it difficult to get ahead of people in lines at games, theaters, or other places. | + | 1.5 | | + | Regular |
| 15 | 24 | I have found it easy to get ahead of people in lines at games, theaters, or other places. | — | | 3.5 | + | |
| Emotional instability | | | | | | | |
| 4 | 25 | I usually feel I am getting ahead in doing things that I would like to do. | + | 4.6 | | + | Nonpo- larity |
| 16 | 26 | I often feel as though I am held back from doing things that I would like to do. | — | 2.7 | | — | |
| 39 | 27 | I find it rather easy to relax and take things easy. | + | 2.0 | | + | Regular |
| 17 | 28 | I find it very difficult to relax and take things easy. | — | | 2.3 | + | |
| Feelings of inadequacy | | | | | | | |
| 1 | 29 | It seems that I manage to do things I would like very much to do. | + | 4.2 | | + | Nonpo- larity |
| 21 | 30 | It seems I'm left out of things I would like very much to be doing. | — | 3.2 | | — | |
| 40 | 31 | I found I can be quite as successful as most of my acquaintances. | + | | 3.6 | — | Reversed polarity |
| 33 | 32 | I found it almost impos- sible to be as successful as most of my acquaint- ances. | — | 4.0 | | — | |
| | | Bal. forward | | 42.2 | 56.3 | | |

TABLE 1 (continued)

| (1) Random | (2) System- atic | (3) Dimensions | (4) + or — | (5) Speed of Agree | (6) of Dis- agree | (7) + or — | (8) Inter- pretive patterns |
|---------------------------------------|------------------------|--|---------------------|--------------------------|----------------------------|---------------------|--------------------------------------|
| <i>Subject A (wholesome, mature)</i> | | | | | | | |
| | | Bal. forward | | 42.2 | 56.3 | | |
| | | <i>Physical defects</i> | | | | | |
| 7 | 33 | I don't care about my physical defects even if they should detract from my appearance. | + | 2.8 | | + | Irregular |
| 34 | 34 | I have been unhappy because of some physical defects that detract from my appearance. | — | | 1.7 | + | |
| 30 | 35 | I am not troubled that people seem to notice certain physical defects I have. | + | | 2.8 | — | Nonpopularity |
| 5 | 36 | I am troubled because people seem to notice certain physical defects I have. | — | | 2.6 | + | |
| | | <i>Nervous manifestations</i> | | | | | |
| 32 | 37 | I have no difficult time sleeping. | + | 1.8 | | + | Irregular |
| 18 | 38 | I have a difficult time sleeping. | — | | 1.6 | + | |
| 24 | 39 | I found I am not tired much of the time. | + | 3.6 | | + | Irregular |
| 23 | 40 | I found I am tired much of the time. | — | | 3.4 | + | |
| | | Total reaction time | | 50.4 | 68.4 | | |
| | | Mean reaction time | | 2.6 | 3.2 | | |
| | | | | 2.9 | | | |
| | | Lengthened reaction time | | 4.4 | and above | | |
| <i>Subject B (neurotic, immature)</i> | | | | | | | |
| | | <i>Close personal relationships</i> | | | | | |
| 36 | 1 | If I am in trouble, there is someone to whom I can go. | + | 1.6 | | + | Regular |
| 2 | 2 | There is no one to whom I can go if I am in trouble. | — | | 8.0 | + | |
| | | Bal. forward | | 1.6 | 6.0 | | |

TABLE 1 (continued)

| (1) Numbers Random | (2) System- atic | (3) Dimensions | (4) + or — | (5) Speed Agree | (6) of Dis- agree | (7) + or — | (8) Inter- pre- tive patterns |
|--------------------------------|------------------------|---|---------------------|-----------------------|----------------------------|---------------------|---|
| Subject B (neurotic, immature) | | | | | | | |
| | | Bal. forward | | 1.6 | 6.0 | | |
| 27 | 3 | I have a number of friends among members of the opposite sex. | + | 2.1 | | + | Irregular |
| 14 | 4 | I have no friends among members of the opposite sex. | — | | 1.3 | + | |
| Interpersonal skills | | | | | | | |
| 20 | 5 | I usually show an interest in the things my acquaintances are doing. | + | 5.3 | | + | Regular |
| 19 | 6 | I rarely show an interest in the things my acquaintances are doing. | — | | 7.6 | + | |
| 28 | 7 | It is easy for me to be friendly with people even when they disagree with me. | + | | 1.5 | — | Reversed polarity |
| 8 | 8 | It is difficult for me to be friendly with people when they disagree with me. | — | 5.5 | | — | |
| Social participation | | | | | | | |
| 11 | 9 | I am a member of a club or other group where people do interesting things together. | + | 1.0 | | + | Regular |
| 9 | 10 | I am not a member of a club or other group where people do interesting things together. | — | | 1.3 | + | |
| 22 | 11 | When I have time I call on my friends. | + | | 1.8 | — | Reversed polarity |
| 29 | 12 | When I have time I don't call on my friends. | — | 2.7 | | — | |
| Satisfying work and recreation | | | | | | | |
| 35 | 13 | I usually look forward with pleasure to the duties of each new day. | + | 8.1 | | + | Irregular |
| 6 | 14 | I rarely look forward with pleasure to the duties of each day. | — | | 1.9 | + | |
| | | Bal. forward | | 26.3 | 21.4 | | |

TABLE 1 (continued)

| (1) Random | (2) System- atic | (3) Dimensions ^a | (4) + or — | (5) Speed Agree | (6) of Dis- agree | (7) + or — | (8) Inter- pre- tive patterns |
|--------------------------------|------------------------|--|---------------------|-----------------------|----------------------------|---------------------|---|
| Subject B (neurotic, immature) | | | | | | | |
| | | Bal. forward | | 26.3 | 21.4 | | |
| 38 | 15 | I have as much time for play and recreation as I should. | + | 1.5 | | + | Regular |
| 37 | 16 | I have less time for play and recreation as I should. | — | | 2.1 | + | |
| Adequate outlook and goals | | | | | | | |
| 26 | 17 | I make plans for carrying on a successful career. | + | 10.5 | | + | Nonpo- larity |
| 3 | 18 | I don't make plans for carrying on a successful career. | — | 8.8 | | — | |
| 12 | 19 | I feel that there is some great plan in the universe and that I have a part in it. | + | | 1.5 | — | Reversed polarity |
| 25 | 20 | I don't feel that there is some great plan in the universe and that I have a part in it. | — | 15.0 | | — | |
| °Behavioral immaturity | | | | | | | |
| 13 | 21 | I find that I get along best when I concern myself about other people's feelings. | + | 2.1 | | + | Regular |
| 10 | 22 | I find I get along best when I don't concern myself about other people's feelings. | — | | 3.1 | + | |
| 31 | 23 | I have found it difficult to get ahead of people in lines at games, theaters, or other places. | + | 1.0 | | + | Irregular |
| 15 | 24 | I have found it easy to get ahead of people in lines at games, theaters, or other places. | — | | 0.8 | + | |
| | | Bal. forward | | 65.2 | 28.9 | | |

TABLE 1 (*continued*)

| (1) Numbers Random System- atic | (2) | (3) Dimensions | (4) + or — | (5) Speed of Agree | (6) of Dis- agree | (7) + or — | (8) Inter- pre- tive patterns |
|--|-----|--|---------------------|--------------------------|----------------------------|---------------------|---|
| <i>Subject B (neurotic, immature)</i> | | | | | | | |
| | | Bal. forward | | 65.2 | 28.9 | | |
| | | <i>Emotional instability</i> | | | | | |
| 4 | 25 | I usually feel I am getting ahead in doing things that I would like to do. | + | | 1.5 | — | } Reversed polarity |
| 16 | 26 | I often feel as though I am held back from doing things that I would like to do. | — | 1.5 | | — | |
| 39 | 27 | I find it rather easy to relax and take things easy. | + | | 1.3 | — | } Reversed polarity |
| 17 | 28 | I find it very difficult to relax and take things easy. | — | 0.8 | | — | |
| | | <i>Feelings of inadequacy</i> | | | | | |
| 1 | 29 | It seems that I manage to do things I would like very much to do. | + | | 1.5 | — | } Reversed polarity |
| 21 | 30 | It seems I'm left out of things I would like to be doing. | — | 2.4 | | — | |
| 40 | 31 | I found I can be quite as successful as most of my acquaintances. | + | 2.5 | | + | } Irregular |
| 33 | 32 | I found it almost impossible to be as successful as most of my acquaintances. | — | | 1.9 | + | |
| | | <i>Physical defects</i> | | | | | |
| 7 | 33 | I don't care about my physical defects even if they should detract from my appearance. | + | | 1.6 | — | } Nonpolarity |
| 34 | 34 | I have been unhappy because of some physical defects that detract from my appearance. | — | | 1.0 | + | |
| | | Bal. forward | | 72.4 | 37.7 | | |

TABLE 1 (continued)

| (1) Numbers Random | (2) System- atic | (3) Dimensions | (4) + or — | (5) Speed of Agree | (6) Dis- agree | (7) + or — | (8) Inter- pre- tive patterns |
|---------------------------------------|------------------------|---|---------------------|--------------------------|----------------------|---------------------|---|
| <i>Subject B (neurotic, immature)</i> | | | | | | | |
| | | Bal. forward | | 72.4 | 37.7 | | |
| 30 | 35 | I am not troubled that people seem to notice certain physical defects I have. | + | | 7.5 | — | Nonpolarity |
| 5 | 36 | I am troubled because people seem to notice certain physical defects I have. | — | | 2.1 | + | |
| | | <i>Nervous manifestations</i> | | | | | |
| 32 | 37 | I have no difficult time sleeping. | + | | 1.0 | — | Reversed polarity |
| 18 | 38 | I have a difficult time sleeping. | — | 0.6 | | — | |
| 24 | 39 | I found I am not tired much of the time. | + | | 1.2 | — | Reversed polarity |
| 23 | 40 | I found I am tired much of the time. | — | 1.1 | | — | |
| | | Total reaction time | | 74.1 | 49.5 | | |
| | | Mean reaction time | | 3.9 | 2.35 | | |
| | | | | 3.13 | | | |
| | | Lengthened reaction time | | 4.72 and above | | | |

Out of 40 polar sets, subject A produced 15 polar reactions. 8 were "regular"; 3, "irregular"; 5, "nonpolar"; 4, "reversed polar." Subject A produced 27 positive indicators *vs.* 13 negative indicators.

Out of 40 polar sets, subject B produced 17 polar reactions. 5 were "regular"; 4, "irregular"; 3, "nonpolar"; 8, "reversed polar." Subject B produced 21 positive indicators *vs.* 19 negative indicators.

In order to conceal the cross-examination arrangement of the test items from the subject, distracter stimulus items, at the rate of one for every two items of the "Random List of Dimensions," are interspersed among the actual test items and are also administered to the subject. As the distracter items form no organic part of the test instrument proper, they are not presented here.

The test is administered by presenting each item to the subject for acceptance or rejection, if felt congenial or uncongenial by the subject, respectively. The subject indicates his choice by depressing an "approving" or "disapproving" button. The reaction time from the moment of presentation

is ascertained by means of a chronoscope. For trained experimenters, testing individual subjects for clinical purposes, a simple stopwatch is adequate.

The approval or disapproval reaction time is then entered in column 3, Table 2 (not illustrated in this paper), labeled "Speed of 'agree' or 'disagree,'" respectively. For the purpose of interpretation, the reaction times are then transferred from these columns of the "Random List of Dimensions" to columns 5 and 6 in the form, "Dimensions by Categories" (Table 1).

1. *The Key to Table 1*

a. *Column 1.* "Random numbers" are the numbers of items within the random order of administration to the subject.

b. *Column 2.* This column contains the "Systematic numbers," that is, the numbers within the systematic order of items in Table 1.

c. *Column 3.* This column lists the stimulus propositions. It was found that negative phrasing of stimulus propositions causes some subjects to take more time for grasping the meaning than is needed for the understanding of the positive counterpart of the statement. Such an effect would, of course, distort the reaction-time scores on the respective cross-examination sets. If lengthened reaction time occurs, it merely should reflect a subject's emotional block on that particular item, not his puzzlement at the grammatical construction. For this reason, negative constructions of stimulus propositions are avoided for subjects of limited education and others who are possibly perturbed by the phrasing of the items.

d. *Column 4.* A *plus* means that the proposition on the left of the plus is "strong" in terms of the underlying motive. If the subject signals his agreement with such an item by depressing the "approval" button, he reveals strength in the motive on which the item has been patterned. By contrast, we find a *minus* on the left of the "weak" items, that is of such as are patterned on the weak pole of the particular motive.

e. *Columns 5 and 6.* These columns contain the reaction times, transferred to them from the "Random List of Dimensions" (Table 2—not reproduced here). They are labeled, "Speed of 'agree' or 'disagree.'"

When all reaction times are posted on the systematic list ("Dimensions by Categories," Table 1), the tester is in a position to survey the whole of the subject's reaction-time patterns emerging from the cross-examinative sets (see following column).

f. *Column 7.* This column, headed "+ or —," is the basis of the "interpretive patterns" in column 8. A *plus* is entered in column 7, if the subject signaled his acceptance of a "strong" item (marked by *plus* in column 4);

or a *minus* is entered if the subject rejected a "weak" item (marked by *minus* in column 4). If he showed, however, agreement with a "weak" item (preceded by *minus* in column 4), or rejected a "strong" item (preceded by *plus* in column 4), a *minus* is entered in column 7.

Thus, two signs, *plus* or *minus*, result from each "cross-examinative set." To be sure, these signs are qualified by the subject's reaction times in making his "agree" or "disagree" decisions. In this way, the following "interpretative patterns," entered in column 8, result.

g. Column 8.

(1). *Regular interpretative patterns.* These indicate a subject's complete wholesomeness on the dimension, underlying the respective cross-examinative set. Thus, subject A is entirely wholesome in the area of "Close Personal Relationships," underlying sets 1, 2 and 3, 4 (see "Dimensions by Categories," Table 1, subject A). Here, the subject manifests congeniality with the "healthy" proposition faster, and uncongeniality with the "unhealthy" item more slowly.

(2). *Irregular interpretative patterns.* These signify the subject's choice—in the present illustration—of a "healthy" stimulus proposition, but the choice is being made with some reluctance, as it were, as expressed by the longer reaction time than for the counter-item. Examples in point are 33, 34 (on "Physical Defects"); and 37-40 (on "Nervous Manifestations") (see, "Dimensions by Categories," Table 1, subject A). In cases such as these, the subject's respective mental health characteristics are somehow "sicklied over," as the lengthened reaction times on the "healthy" propositions bespeak. An "inquiry" into these particular aspects by way of directive interview reveal the nature of the emotional disturbance.

(3). *Nonpolar interpretative patterns.* These emerge when the subject accepts both the "strong," and "healthy," etc., and the "weak," "unhealthy," etc., motive pole, for example 9, 10 and 25, 26 in the present illustration (see "Dimensions by Categories," Table 1, subject A). Nonpolar interpretative patterns reflect a subject's vacillation or ambivalence regarding the underlying dimension. As validation studies showed, even in nonpolar interpretative patterns, the item having the shorter reaction time seems to reflect the subject's personal disposition more closely than does the item with the longer reaction time.

(4). *Reversed polarity interpretative patterns.* These reveal the subject's choice of the "weak," "unhealthy," etc., motive pole and his rejection of its "strong," "healthy," etc., counterpart. A case in point in our illustration is

found in items 7, 8 (on "Interpersonal Skills," in subject A) ("Dimensions by Categories," Table 1).

On surveying the reaction times in columns 5 and 6 (Table 1), one will notice some conspicuously lengthened times, for example, 25 and 28 in subject A. Lengthened reaction times are in bold print. In the present test, reaction times are regarded as excessively lengthened if exceeding the mean of all of the subject's reaction times by 50 per cent or more.

Conspicuously lengthened reaction times reflect emotional complexes underlying the dimensions on which the diagnostic patterns were patterned. Inferences from excessive reaction times taken by subjects in supplying reaction words to stimulus words are, of course, one of the basic indicators of complexes in the classical C. G. Jung Word Association Test.

At the end of the form ("Dimensions by Categories" Table 1), the reader will find several quantitative formulas, for example, the ratio between the positive and the negative indicators entered in column 7. A comparison between the ratio 27:13 in subject A, the wholesome criterion person, to 21:19 in subject B, the neurotic criterion person, is significant.

From the "interpretive patterns" (column 7, Table 1), the following interpretations of subject A's and subject B's test protocols were derived. Before presenting them, it must be pointed out that the "technique of timed cross-examination" embodies the signal advantage of not depending on a highly trained interpretive skill of the diagnostician.

The subject's concurrent voluntary and involuntary responses can either reflect wholesome or emotionally disturbed dimensions. "Wholesome dimensions" are revealed by "regular interpretative patterns," as defined above. "Emotionally disturbed dimensions," however, are reflected by all other patterns, described above. Deviations from "regular patterns" signify the subject's involuntary reactions, coupled with the concurrent voluntary reactions. Thus, the subject's involuntary reactions provide "projective clues" to his concurrent voluntary responses.

Thanks to this setup, the interpretation of the subject's responses does not depend on "interpretative intuition." By following the "interpretive patterns" of column 8 in Table 1, the tester obtains the interpretations of each cross-examinative set separately. This amounts to no more than to modifying the given text of the stimulus propositions according to the special meanings of the "interpretive patterns." Any intelligent interpreter can meet this largely semantic requirement without extensive training.

The interpretation is largely evolved through the "lie-detector device" which "timed cross-examination" affords. Deceptive attempts of the subject,

or his self-delusions, far from detracting from the objectivity of the procedure, are detected through the subject's speed of responses, qualifying the responses. Discrepancies between the concurrent voluntary and involuntary responses reveal objectively where those discrepancies occur.

C. INTERPRETATION OF TWO CONTRASTING-CRITERION SUBJECTS' MENTAL HEALTH PATTERNS

The reader is invited to follow the "interpretive patterns" in column 8 of Table 1 and the respective stimulus sets in column 3. In the following interpretations, the sets are indicated by their "systematic numbers," listed in column 2; the subjects' areas of emotional disturbance are italicized, but are not to be confused with the italicized side headings.

1. *Subject A*

a. *Close personal relationships.* [1, 2]: A is thoroughly mature in her personal relationships.

b. *Interpersonal skills.* [5, 6]: She shows interest in the things others are doing; but see below [21, 22]. [7, 8]: She does not pretend to be friendly with people who disagree with her.

c. *Social participation.* [11, 12]: She seeks individual friendships; [9, 10]: but has mixed feelings as to participation in clubs and other groups.

d. *Satisfying work and recreation.* [13, 14]: *She does not look forward with pleasure to the duties of each day;* [15, 16]: but seems to enjoy her time off duty.

e. *Adequate outlook and goals.* [17, 18]: She claims not to care about career; [19, 20]: but finds her life, and life altogether, meaningful.

f. *Behavioral immaturity.* [21, 22]: Her attitude toward other people is ambivalent. [23, 24]: She hates to "use her elbows."

g. *Emotional instability.* [25, 26]: *Regarding her aspirations she reveals an emotional block.* Notice the lengthened reaction time on item 25. [27, 28]: *Her claim to find it rather easy to relax and to take things easy must be taken with a grain of salt.* See lengthened reaction time on item 28.

h. *Feelings of inadequacy.* [29, 32]: *The preceding indication is supported by her feelings of inadequacy.*

i. *Physical defects.* [33, 34]: She does not seem to suffer from physical defects; [35, 36]: *yet is not free from self-consciousness regarding physical appearance.*

j. *Nervous manifestations.* [37, 40]: *She would like to get more rest and seems to live under pressure, without, however, succumbing to it.*

The "inquiry" of the subject and pooled judgments by her acquaintances bore out the above results obtained by "timed cross-examination." Most of the detected emotional blocks seem to be explained by the fact that the subject is a very gifted painter, but is compelled by circumstances to do library work for a living.

The preceding interpretation of a wholesome, mature-criterion person's "timed cross-examination" protocol is followed by that of a neurotic-criterion person. Note again the italicized areas of emotional disturbance.

2. Subject B (Neurotic)

a. *Close relationships.* [1, 2]: *If he is in trouble, there is someone to whom to go. Notice, however, the lengthened reaction time on item 2.* [3, 4]: *He has friends among the opposite sex, but this relationship creates a problem for him.*

b. *Interpersonal skills.* [5, 6]: *He usually takes interest in the things his acquaintances are doing. Yet lengthened reaction times betray underlying emotional blocks.* [7, 8]: *It is not easy for him to be friendly with other people when they disagree with him. The lengthened reaction time on item 8 reveals the emotional disturbance—probably guilt feelings—underlying this situation.*

c. *Social participation.* [9, 10]: *He likes to join groups; [11, 12]: but does not seek individual friends.*

d. *Satisfying work and recreation.* [13, 14]: *He claims to look forward with pleasure to the duties of each new day. But notice the lengthened reaction time on item 13.* [15, 16]: *His enjoyment of pleasure and recreation is genuine.*

e. *Adequate outlook and goals.* [17, 18]: *Subject has obviously an emotional block in connection with his future career. Judging from the shorter reaction time on item 18, that is, the alternative of not making plans, this seems to be his present unavowed attitude. Perhaps he is shrinking from responsibility and prefers the sheltered life in the hospital. This became clear in the "inquiry" after the test. The "inquiry" revealed that he is seeing his wife regularly on weekends, with the hospital's permission. She holds a gainful position.* [19, 20]: *This reaction seems to be crucial in the total assessment: Subject rejected the meaningfulness of his existence very rapidly: see item 19. Item 20 corroborates this attitude. It has an extremely long reaction time.*

f. *Behavioral immaturity.* [21, 22]: *Subject is interested in other people; yet the "irregular interpretive pattern," resulting from [23, 24], reveals an*

underlying emotional complex, perhaps repressed aggressive feelings toward other people.

g. Emotional stability. [25, 26]: This set reveals subject's inferiority complex and defeatism. [27, 28]: Subject finds it very difficult to relax and to take things easy.

h. Feelings of inadequacy. [29, 30]: This set reflects strongly subject's feelings of inadequacy, already found in [25, 26]. [31, 32]: Subject's boast of standing his ground in competition with others is at striking variance with the preceding set. Subject's case history states that before his nervous breakdown he was a competent employee, with executive capabilities. It is, however, doubtful that he could live up to his boast in his present frame of mind.

i. Physical defects. [33, 34]: Subject's admission in item 33 of being disturbed by physical defects is opposed by his denial of being so disturbed in item 34. Having shorter reaction time than item 33, item 34 seems to be more natural to the subject. This is further confirmed by [35, 36]. Though ambivalent as to admission or denial of worry about physical defects, item 36, reflecting admission, has the shorter reaction time.

j. Nervous manifestations. [37-40]: Subject has a difficult time sleeping and is tired much of the time.

In retrospect we find that this neurotic subject is emotionally disturbed in virtually each of the listed mental health aspects. There are only three wholesome areas, free from emotional blocks: subject's social participation [9, 10]; his interest in other peoples' feelings [21, 22]; and his enjoyment of recreation [15, 16].

These findings were confirmed in the "inquiry." When asked what he enjoys most in the hospital, he mentioned participation in group activities where he can take a hand. But as to future plans when released, he was evasive (see *Adequate outlook and goals*, above).

D. TOWARD THE AUTOMATION OF PERSONALITY ASSESSMENT THROUGH THE TECHNIQUE OF "TIMED CROSS-EXAMINATION"

The technique used for the testing of the preceding and other contrasting criterion persons was rather primitive: Each stimulus item was read at an even speed to the subject by the experimenter; the subject's voluntary reactions consisted merely in saying "yes" or "no" according to his acceptance or rejection of the respective item; and the involuntary speed of the reactions was measured by a simple stopwatch, started after finishing the reading of

each item and stopped at the moment of the subject's response of "yes" or "no."

The reader must have realized the efficacy of the technique. The test scores are obtained automatically: "agree" or "disagree," concurrent with the speed of "agree" or "disagree." Also, the test scores are objective, not projective. So is the interpretation. It is based on the "interpretive patterns" (column 8, Table 1), which in turn are derived from the objective test scores. Actually, projective effects are here achieved by objective means, since the objectively obtained reaction patterns may reveal subconscious areas of emotional disturbance, or their absence. Thus, no projective guesswork is involved in the interpretation of "timed cross-examination" protocols.

A limited degree of automation can be achieved by a setup of equipment such as the one described below.

For *visual presentation* an apparatus such as Model 340 Card Master, manufactured by Hunter, Manufacturing Company, Iowa City, Iowa, is suitable. This is a cycle card display machine by automatic or manual operation. A relay in the unit starts an electronic timer whenever the shutter opens, and stops it whenever the subject depresses the "agree" or "disagree" button.

If *acoustic* instead of visual presentation is chosen—perhaps because of its better suitability for certain types of subjects—a dual-channel tape recorder can be employed. As soon as the tape recorder starts "talking," a noise-operated relay will detect the noise and start the electronic timer, keeping it running until the subject depresses the "agree" or "disagree" button.

The reaction times on the subject's acceptances or rejections of stimulus propositions could, of course, be entered by the tester in the respective columns of the "Random List of Dimensions." This time-consuming procedure could, however, be avoided by the use of a print-out-counter, which would record the subject's reaction times in two separate columns. The sum totals of each column, as well as their arithmetic means and the "long reaction times," that is, the ones conspicuously above average, should be obtained by a computer. Present day computer technique should make it possible to produce the "interpretive patterns" of column 7 (Table 1) automatically.

Model 4600 Print-out Counter, manufactured by Grason-Stadler, West Concord, Massachusetts, is automated to the point of storing the reaction times and of printing them out in different separate columns.

The hardest stage amenable to automation is the interpretation of the

subject's responses. It will not have escaped the reader's notice that the wording of the interpretation in the illustrative cases of this paper follows closely the wording of the stimulus items, modifying it, however, according to the "interpretive patterns" of column 8, "Dimensions by Categories" (Table 1).

The automation of the interpretation would necessitate an especially adapted computer into which all possible modifications of the stimulus texts by all possible "interpretive patterns" would be fed. Then, by tapping on a specific "interpretative pattern" obtained in a given test, the respective interpretation would be printed out by the computer. Manufacturers specialized along this line are capable of constructing equipment of this sort according to the psychologist's specifications.

E. CONCLUSION

In the present stage of research into the "technique of timed cross-examination" complete automation would be premature. What matters in the first place is to perform on a larger scale than heretofore a test of significance of the procedure, by testing contrasting criterion groups and by analyzing the results statistically. If once this requirement be satisfactorily met, the "technique of timed cross-examination" could be adapted for the testing purposes of a great variety of users, such as psychological clinics, schools and universities, business and industrial enterprises, as well as government agencies.

It is to be hoped that the "technique of timed cross-examination," since it is technically based on the simple action of the subject's pressing choice buttons and on his concurrent recording of the speeds of his choices, bids fair to make possible a far-reaching automation of personality assessment by a technique yielding projective effects through objective means.

REFERENCES

1. BRUNER, J. S. Social psychology and perception. In Maccoby, E. E., *et al.* (Eds.), *Readings in Social Psychology*. New York: Holt, 1958.
2. BRUNER, J. S., & POSTMAN, L. On the perception of incongruity. *J. Personal.*, 1949, **18**, 206-223.
3. BUYTENDIJK, F. J. *Traité de Psychologie Animale*. Paris: Presses Univ., 1952.
4. GILBERT, A. R. Projective cross-examination. *J. of Psychol.*, 1956, **42**, 105-132.
5. ———. Self-validation in projective testing. *J. of Psychol.*, 1958, **46**, 203-209.
6. ———. Personality appraisal through "cross-examination." *Psychol. Rep.*, 1960, **6**, 184.
7. ———. The other person: How we 'intend' it. Proceedings of the XVI Internat. Congr. of Psychol., Bonn, 1960, 566/7 (Abstract).

8. ———. The other person: How we 'intend' it. *J. of Psychol.*, 1961, **51**, 247-262.
9. HEIDER, F. *The Psychology in Interpersonal Perception*. New York: Wiley, 1958.
10. JUNG, C. G. The association method. *Amer. J. Psychol.*, 1910, **21**, 219-269.
11. ———. *Studies in Word Associations*. New York: Moffat, Yard, 1919.
12. LEVINE, J. M., & MURPHY, G. The learning and forgetting of controversial material. *J. Abn. & Soc. Psychol.*, 1943, **38**, 507-515.
13. TAGIURI, R., BRUNER, J. S., & BLAKE, R. On the relations between feeling and perception of feeling among members of small groups. In Maccoby, A. A., et al. (Eds.), *Reading in Social Psychology* (3rd ed.) New York: Holt, 1958.

Department of Psychology
West Virginia Wesleyan College
Buckhannon, West Virginia

SELF-DISCLOSURE AND EXPRESSED SELF-ESTEEM, SOCIAL DISTANCE AND AREAS OF THE SELF REVEALED*¹

Department of Psychology, Marymount Manhattan College

MAUREEN P. FITZGERALD

A. INTRODUCTION

A fundamental premise in interpersonal relations is that "People need to receive from, and to give to" (5, p. 1). Yet, what one gives, or is willing to give, how much, and to whom are aspects of the social distance dimension that are relatively unexplored. Jourard and Lasakow (4) have termed this area of social distance "self-disclosure," defining it as the process of making the self known to another person, the readiness to confide personal information to others, or even to misrepresent the self to others. Jourard (2) suggests that the accurate portrayal of the self to others is an identifying criterion of healthy personality. The amount of information one person is willing to disclose to another appears to be an index of the "closeness" of the relationship. Subjects tend to have the same amount disclosed to them as they in turn disclose to others (2, 3).

In a study of social distance Jourard and Lasakow (4) developed a 60-item self-disclosure questionnaire covering six areas: attitudes and opinions, tastes and interests, work or studies, money, personality, and body. In this instrument the subject indicated the extent to which he *has* made himself known to another called a "target person"—a person to whom information about the self is communicated. The data indicated that whites disclosed more than Negroes, and females disclosed more than males. More was disclosed to the preferred (better-liked) parent, and more to the mother than to the father.

In seeking to determine basic factors underlying self-disclosure, it was felt that self-esteem might influence the freedom with which one would disclose the self to another. Thus, those with high self-esteem feeling more secure about their attitudes and finding less need to gain approval or support

* Received in the Editorial Office on June 12, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ Submitted in partial fulfillment of the Ph.D. requirements in the Psychology Department, Fordham University. The writer wishes to express sincere appreciation to Professor Joseph Kubis for directing this investigation.

from others, would find it less necessary to volunteer information about themselves. In similar fashion, those expressing low self-esteem may feel the need for self-disclosure to gain attention. It is therefore hypothesized that subjects with expressed high self-esteem will disclose less than those with expressed low self-esteem.

Regardless of the amount of expressed self-esteem, it would seem likely that the greater the degree of liking for another, i.e., the closer one felt to another, the more would be disclosed to this other. One should expect individuals to disclose more of the self to one liked best, less to an "average"—one who is neither a close friend nor disliked—and still less to one liked least.

The willingness to disclose information about the self is probably affected by the nature of the self-information. Information pertaining to the more "public" aspects of the self such as interests and attitudes, work or study, is probably revealed more readily while information pertaining to the "private" self, more ego-involved material, is less readily revealed.

Using a group of 300 college women as subjects, the following three hypotheses were tested: (*a*) subjects with expressed high self-esteem will disclose less than those with expressed low self-esteem; (*b*) there will be a greater amount disclosed to a girl in the class liked best, a lesser amount to an "average" girl (one midway on the social distance continuum), and a still lesser amount to the girl in the class liked least; and (*c*) there will be more disclosed in the "public" areas of tastes and interests, attitudes and opinions, and work and study than in the "private" areas of money, personality, and body.

B. METHOD

1. *Subjects*

The 300 subjects were unmarried, white, female day-students at a small denominational college for women in New York. There were 75 each of freshmen, sophomores, juniors, and seniors, most coming from middle-class socioeconomic homes in the metropolitan area. It is the policy of the college to plan a series of events to integrate each student actively into college life. Thus, by December of her freshman year, each student knows every other member of her class by name and by sight, and has had some interaction with each of them.

2. *Instruments Used*

For an index of expressed self-esteem, each subject filled out a self-rating scale on which the subject was requested to "describe yourself as you ordi-

narily think about yourself" (1). This self-rating scale consisted of 20 bipolar pairs of adjectives—a trait and its opposite. Each bipolar trait had a six-point scale. Each point was assigned a numerical value ranging from six to one. The positive or favorable pole for each pair of items was assigned a value of six, while the negative or less favorable pole was assigned a value of one. The four units between these two poles were assigned values of five, four, three, and two. The Self-Esteem ratings on each of the 20 items were added for one score. The highest score possible was 120. The upper, middle, and lowest 20 per cent in each college class were selected on the basis of the total score, and constituted the High, Median, and Low Self-Esteem groups respectively.

a. Self-disclosure. A modified form of Jourard and Lasakow's (4) 60-item questionnaire was filled out by each of the 300 subjects for the amount of information that had been disclosed to three different "target persons." Since all of the subjects used in the present study were unmarried females, the following two modifications were made: In the area of personality, statement three was replaced by: "my attitudes towards men and the problems involved in this area." In the area of body, statement 10 was replaced by: "my feelings about and attitudes towards intimate feminine problems." The categories, directions, scoring procedure, etc., are that of Jourard and Lasakow (4). Each subject filled out this questionnaire three times: (a) "how much you have revealed to the girl in the class whom you like best;" (b) "the girl in the class you like least;" and (c) "a specific girl in the class who is not a friend of yours nor the girl whom you like least. Think of a girl who is a typical or average girl in your class. Indicate how much you have talked with her about this item."

The following five indices were obtained for each subject: (a) Self-Esteem; (b) Self-Disclosure to the girl in the class liked best, (c) to the girl liked least, (d) and to the average girl in the class; and (e) Total Self-Disclosure from a sum of indices (b), (c) and (d).

C. RESULTS

Split half odd-even reliability coefficients were computed for each of the six areas of the questionnaire for each of the four college classes. The Spearman-Brown correction formula was applied and the resulting reliability coefficients ranged from .78 to .99 with 20 of the 24 having values over .90. Hence, the self-disclosure questionnaire was considered a reliable instrument.

Total-score means and standard deviations for the various college groups appear in Table 1. In order to ascertain if groups varying in the amounts of

self-esteem would differentiate themselves in the amount of self-disclosure, the data were subjected to an analysis of variance. The variances proved to be homogeneous, but none of the *F* ratios reached significance.

An inspection of the Mean scores (Table 1) reveals that the least amount

TABLE 1
MEANS AND STANDARD DEVIATIONS FOR TOTAL SELF-DISCLOSURE TO THREE TARGET PERSONS FOR THREE SELF-ESTEEM GROUPS FOR FOUR COLLEGE CLASSES

| Groups | Means | SD |
|-------------------------------|--------|-------|
| Low Self-Esteem Sophomores | 129.07 | 40.07 |
| High Self-Esteem Seniors | 117.47 | 31.70 |
| High Self-Esteem Juniors | 116.47 | 45.30 |
| Median Self-Esteem Seniors | 115.33 | 27.48 |
| High Self-Esteem Sophomores | 115.27 | 36.50 |
| Low Self-Esteem Juniors | 106.20 | 33.98 |
| Median Self-Esteem Juniors | 104.13 | 34.25 |
| High Self-Esteem Freshmen | 104.00 | 43.70 |
| Median Self-Esteem Juniors | 102.93 | 27.00 |
| Low Self-Esteem Seniors | 99.40 | 33.35 |
| Median Self-Esteem Sophomores | 93.98 | 32.39 |
| Low Self-Esteem Freshmen | 93.93 | 31.20 |

disclosed (93.93) was made by the Low Self-Esteem freshmen, and the third least amount disclosed (99.40) was by the Low Self-Esteem seniors, while the second and third greatest amounts disclosed was that by two High Self-Esteem groups—seniors (117.47) and juniors (116.47). A further inspection of this Table indicates that there is no patterning of the Self-Esteem groups with reference to the amount of self-disclosure.

The second aspect of self-disclosure investigated was that of the effect of the "closeness" felt to the target person by the subject—how much the subject liked the target person—on the amount disclosed. It would seem likely that regardless of the amount of self-esteem, the greater the degree of liking the subject has for another, the more would be disclosed to this person. From Table 2, it can be seen that the mean amount of information disclosed to the girl liked best was 68.47; to the average girl less

TABLE 2
MEANS AND STANDARD DEVIATIONS FOR SELF-DISCLOSURE TO THREE TARGET PERSONS—GIRL LIKED BEST, AVERAGE GIRL AND GIRL LIKED LEAST

| Target person | Means | SD |
|-------------------------|-------|-------|
| Girl liked best | 68.47 | 20.70 |
| Typical or average girl | 27.02 | 14.54 |
| Girl liked least | 13.86 | 12.22 |

than half of this (27.02); and to the girl liked least, half of that again (13.86).

A line of regression was computed (6) to test for the linearity of the relationship between the amount of disclosure and the degree of closeness to the target person. Table 3 indicates a regression of 27.25 units of self-disclosure from one degree of liking to another. Although some of the

TABLE 3
SELF-DISCLOSURE REGRESSION ON DEGREES OF LIKING FOR TARGET PERSONS

| Degree of liking | Liked best | Average or typical | Liked least |
|------------------|----------------------------------|--------------------|-------------|
| Sum | 20,238 | 7,994 | 4,107 |
| | +1 | 0 | -1 |
| | Sum of squares, degree of liking | | 478,866.52 |
| | Linear regression | | 439,542.50 |
| | Deviation from regression | | 39,324.02 |
| | Regression per unit of liking | | 27.25 |

regression is accounted for by linearity, there is appreciable deviation from linearity. The typical or average girl in the class does not fall midway between the girl liked best and the girl liked least with reference to the amount disclosed to her, but, rather, falls nearer to the girl liked least. Thus the second hypothesis, that there would be a greater amount disclosed to the girl liked best, a lesser amount to the "average" girl, and a still lesser amount to the girl liked least, has been supported by the data.

Data of Jourard and Lasakow (4) on self-disclosure to four different target persons—mother, father, male friend, and female friend—are presented

*TABLE 4
MEANS AND STANDARD DEVIATIONS FOR SELF-DISCLOSURE TO FOUR TARGET PERSONS—
MOTHER, FATHER, MALE FRIEND, FEMALE FRIEND—AS REPORTED BY JOURARD AND
LASAKOW

| Target person | Means | SD |
|---------------|-------|-------|
| Mother | 72.30 | 19.50 |
| Father | 51.70 | 24.13 |
| Male friend | 55.18 | 22.43 |
| Female friend | 56.58 | 27.20 |

in Table 4 based on a group composed of males and females, Negroes and whites. As was stated earlier, females disclose more than males, and whites disclose more than Negroes. Therefore, in a comparison of the results of Tables 2 and 4 the fact that different target persons were used and that

the composition of the groups also differed, must be borne in mind. Yet it is interesting to note that more was disclosed to the mother (72.30) and the next greatest amount to the girl liked best (68.47). The two target persons to whom the least was disclosed were the "average" girl (27.02) and the girl liked least (13.83). Thus, in both studies the greater the feeling of closeness to the target person—the more the target person was liked—the more was disclosed to that person.

Since the willingness to disclose information about the self is also probably affected by the nature of the information, i.e., pertaining to the "public" self (less ego-involved material), or "private" self (more ego-involved material), *t* ratios were computed for each college class for the mean amounts revealed on the three more "public" areas as compared to the three more "private" areas of the self. As can be seen in Table 5, all the *t* ratios were significant at the .01 level. Significantly more was revealed

TABLE 5
t-RATIOS OF THE DIFFERENCES IN THE AMOUNT OF SELF-DISCLOSURE OF THE "PUBLIC" SELF AND THE "PRIVATE" SELF FOR FOUR COLLEGE CLASSES

| Class | Disclosure means | | <i>t</i> -ratios |
|------------|------------------|--------------|------------------|
| | Public self | Private self | |
| Freshmen | 65.93 | 36.64 | 8.44** |
| Sophomores | 69.53 | 41.65 | 7.99** |
| Juniors | 68.31 | 40.44 | 7.87** |
| Seniors | 73.10 | 41.40 | 10.47** |

** Significant at the .01 level.

by all the classes in the areas of the "public" self—tastes and interests, attitudes and opinions, and work or study—than in the more "private" areas of the self—money, personality, and body. Thus, these data support the original hypothesis.

D. DISCUSSION

This investigation has shown that self-disclosure reflects social distance in that one reveals more to the girl liked best than to the "average" girl, and more to the latter than to one liked least. Hence, self-disclosure can be used as an index of social distance. Jourard (2), in a limited study using eight subjects, points to a norm or standard that individuals seem to have pertaining to how much should be revealed about the self. Revealing too little or too much seems to be equally unacceptable. As a friendship is developing, the rate at which information about the self is revealed may be of great importance. Too great an acceleration or deceleration of the rate may retard

or sever the developing friendship. The importance of self-disclosure is reflected in the almost uniform tendency to reveal only a sparse amount to one liked least.

There seems to be some selective factor governing the different kinds of self-information disclosed and also to whom it is disclosed. The significant difference favoring greater amounts of disclosure about more "public" as compared with more "private" aspects of the personality is both reliable and uniform for all four college classes. Subjects are willing to reveal their preferences for particular political ideologies, and their conception of what constitutes suitable dress, but are less willing, or even unwilling, to reveal their attitudes toward their own bodies. The former are probably discernable to an astute observer without being so informed by the subject. These topics constitute what is generally termed "social conversation." In contrast, information pertaining to the more private areas of the personality, e.g., money, is considered in "poor taste" if given forth too freely.

It had been hypothesized that those expressing high self-esteem, feeling more secure about their attitudes would find it less necessary to volunteer information about themselves, while those expressing low self-esteem would feel the need for self-disclosure to gain attention and/or support. The amount of self-esteem alone does not significantly affect the amount disclosed about the self. Two persons differing markedly in their amount of self-esteem could reveal the same amount of information about the self, but for different reasons. An individual of high self-esteem could reveal the information because she considers her attitudes and opinions worthwhile. One of low self-esteem could reveal the same information to reassure herself, or to test its worthiness via its acceptance or approval by others. It could be possible that some subjects would express high self-esteem due to insecurity or pathology, and may feel compelled to self-disclosure. On the other hand, the withdrawn isolate expressing low self-esteem would hesitate to talk about himself and his feelings of unworthiness. Self-esteem may interact with another or other factors not isolated or not included in this study.

E. SUMMARY

For 300 women college students, indices of expressed self-esteem and self-disclosure to three target persons—a girl liked best, an "average" girl (neither a close friend nor one liked least), and a girl liked least—were obtained. Comparing these indices, it was found that the amount of expressed self-esteem alone does not significantly affect the amount disclosed about the self. Self-disclosure as a dimension of social distance is clearly able to dis-

tinguish how close one feels to another. One discloses significantly more to a girl liked best, significantly less to an average girl, and very little to one liked least. There are also areas of the personality that are more freely disclosed while other areas are disclosed more selectively or not at all. Although "people need people to receive from and to give to" (5), there seem to be some fairly uniform restrictions as to what, how much, and to whom one gives information about the self.

REFERENCES

1. FIEDLER, F., HUTCHINS, E., & DODGE, J. Quasi-therapeutic relations in small college and military groups. *Psychol. Monog.*, 1959, **73** No. 3 (Whole No. 473).
2. JOURARD, S. M. Self-disclosure and other cathexis. *J. Abn. & Soc. Psychol.*, 1959, **59**, 428-431.
3. ———. Self-disclosure scores and grades in nursing college. *J. Appl. Psychol.*, 1961, **45**, 244-247.
4. JOURARD, S. M., & LASAKOW, P. Some factors in self-disclosure. *J. Abn. & Soc. Psychol.*, 1958, **58**, 91-98.
5. SCHUTZ, W. FIRO. New York: Rinehart, 1958.
6. SNEDECOR, G. Statistical Methods. Ames: Iowa State College, 1956.

Department of Psychology
Marymount Manhattan College
221 East 71 Street
New York 21, New York

THE EFFECT OF OBSERVER REDUNDANCY ON DISPLAY MONITORING EQUIPMENT*

Sylvania Electronic Systems, Waltham, Massachusetts

DONALD W. BROWN

A. INTRODUCTION

Much research time and money has been expended over the past two decades in attempting to design displays in such a way that a human is capable of monitoring them efficiently. The basic problem underlying this work is that while systems are becoming more complex and, hence, requiring more complex types of displays, the operating characteristics of humans have been essentially fixed for many thousands of years and there is no reason to suppose that these characteristics will change in the foreseeable future. Given this problem, the approach has been to adapt displays (and systems) to the operator's capabilities.

Two parameters of the operator's response are of prime importance, namely, accuracy and speed. It has been suggested by Williams (7) that the chance of detecting a near-threshold signal is greater for two or more independent observers (assuming approximately equal ability) than for one observer. If this hypothesis can be verified, it provides a method for maintaining a high level of accuracy without sacrificing amount or complexity of displayed information.

It is clear that the probability of occurrence of one of N independent events (P_n) is given by:

$$P_n = 1 - (1 - p_1)(1 - p_2) \dots (1 - p_n). \quad [1]$$

Thus, if p_1, \dots, p_n are the probabilities that 1, \dots , n observers will detect a given signal, the accuracy with which that signal is detected may be improved by adding observers, the higher the probabilities of detected individual observers, the fewer observers need be added to attain a given level of accuracy.

This hypothesis has been tested by Schafer (6) using a sonar detection task. He found that the addition of a second observer of approximately equal ability (average thresholds within 1 db) increased the detection probability by 11 to 20 per cent, depending on the specific task, while a third observer

* Received in the Editorial Office on June 13, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

increased the probability by six to 15 per cent over that realized with two observers. The increases in detection probability were, however, less than those predicted by equation [1]. This deviation from predicted results was explained by Schafer as probably being due to correlation between observers' responses. It should be noted that subjects were instructed that a response in the absence of the stimulus would be counted against them with twice the value of a correct response.

The study reported here was designed to determine the effect on accuracy and speed of performance of task complexity and observer redundancy. Observer redundancy refers to two observers responding as a team to all stimuli as opposed to the nonredundant condition in which each observer responded to half of the stimuli. Unlike Shafer's (6) study, the instructions were designed to provide an equal avoidance set for both positive (false alarm) and negative (missed signal) errors.

The task used in the present investigation consisted of:

1. A display composed of a matrix of identical lights. When certain combinations of lights (critical stimuli) were illuminated, a response was required.
2. All critical stimuli required the same discrete response.
3. The responding element was two human observers.
4. The operators were instructed to respond as quickly as possible to critical stimuli.
5. The number of critical stimuli was constant.

This results in the number of stimuli per observer in the nonredundant condition being one-half the number per observer in the redundant condition.

Assuming the probabilities of positive and negative errors to be equal, the following general hypotheses were derived.

1. Redundant operation results in longer reaction times (RT's) than nonredundant operation. This hypothesis is supported by data [Baxter (2); Saltzman and Garner (5)] showing increases in task complexity resulting in increases in RT. It is hypothesized that this increase in RT due to increased task complexity in the redundant situation will be of greater magnitude than any decrease in RT obtained in the redundant situation as a function of taking the best RT of the two operators.

2. Overall reliability of performance is unaffected by redundancy. The redundant condition results in fewer negative errors than the nonredundant condition; however, this result is reversed with regard to positive errors.

The above general hypotheses, when applied to the present study, resulted in the following specific hypotheses:

1. Nonredundant organization produces shorter RT's than redundant organization.

2. RT's increase as stimulus complexity increases.

3. Redundant organization results in fewer negative errors than nonredundant organization.

4. The number of negative errors increases as stimulus complexity increases and this increase is greater for the nonredundant condition (i.e., there is a significant "redundancy by complexity" interaction).

5. Nonredundant organization results in fewer positive errors (a positive error by either group member in the redundant situation counted as an error, but only one error was possible on any given trial).

6. The number of positive errors increases as stimulus complexity increases and this increase is greater for the redundant condition (i.e., there is a significant "redundancy by complexity" interaction).

B. METHOD

1. *Apparatus*

The apparatus consisted of a matrix of numbered opaque white circles one inch in diameter on a black background which faced the subjects. Each circle was illuminated by a No. 46 pilot light. This size bulb in an ambient illumination of 3 footcandles provided a stimulus well above threshold. The distance between adjacent circles was two inches. Each of the lights was controlled by a toggle switch, so that any combination of lights could be programmed. An additional toggle switch activated the lights which had been programmed as well as a Standard Electric Timer which was graduated in .01 seconds. Two telegraph keys—18 inches apart—were used for responses. A 3 × 5 inch card placed on the table in front of each subject (S) listed the "Critical" light combinations for a given session.

2. *Subjects*

Six groups of two Ss each were comprised of male employees of the Applied Research Laboratory.

3. *Procedure*

All groups of Ss were run under all six combinations of redundancy (redundant and nonredundant) and stimulus complexity (two, three, and four lights per stimulus). The order of occurrence of the six conditions was counterbalanced among groups, and each group received one condition per day.

Before a session was begun, the *Ss* were given 50 practice trials in which they were instructed to depress the telegraph key when the experimenter said "ready" and to release it as quickly as possible at the onset of the number one light. This period also allowed for adaptation to the ambient illumination.

Twelve RT's to critical stimuli were obtained for each pair of *Ss* for each of the six experimental conditions. Four different critical stimuli were used in each condition with the position of the critical stimuli randomized within each experimental condition and group of *Ss*. To allow for the possibility of negative errors, each random series was made up of 64 stimuli of which 16 were "critical"; however, the session ended as soon as the *Ss* had made responses to 12 critical stimuli. The *Ss* were instructed to depress the response key when the experimenter said "ready," and to release it as quickly as possible if a critical stimulus occurred. If the stimulus was non-critical, they were to release the key when the lights were extinguished. The stimulus duration was three seconds and the interval between the "ready" signal and the stimulus onset (foreperiod) varied from 1-3 seconds.

Reaction time, and positive and negative errors were recorded for each session.

C. RESULTS

The Analysis of Variance performed on the RT data is presented in Table 1.

Table 2 summarizes the Analysis of Variance performed on the positive error data.

TABLE 1
SUMMARY OF ANALYSIS OF VARIANCE PERFORMED ON REACTION TIME DATA

| Source of variance | <i>df</i> | <i>M Sq.</i> | <i>F</i> |
|--|-----------|--------------|--------------------------------|
| Redundancy | 1 | 22,605 | 26.470 at 1 + 396 <i>df</i> ** |
| Task complexity | 2 | 8,510 | 9.965 at 2 + 396 <i>df</i> ** |
| Groups | 5 | 2,656 | 3.110 at 5 + 396 <i>df</i> ** |
| Redundancy \times complexity | 2 | 2,557 | 2.994 at 2 + 396 <i>df</i> |
| Redundancy \times groups | 5 | 2,241 | 2.624 at 5 + 396 <i>df</i> ** |
| Complexity \times groups | 10 | 1,631 | 1.910 at 10 + 396 <i>df</i> * |
| Redundancy \times complexity \times groups | 10 | 3,259 | 3.816 at 10 + 396 <i>df</i> ** |
| Within cells (error term) | 396 | 854 | |

** Significant at .01 level.

* Significant at .05 level.

Since there were only seven negative errors in the entire experiment, no analysis could be performed on the data.

The direction of the RT effects revealed by the analysis are shown in Table 3, while the mean positive error scores for the six conditions are shown in Table 4.

TABLE 2
SUMMARY OF ANALYSIS OF VARIANCE PERFORMED ON ERROR DATA

| Source of variance | <i>df</i> | <i>M Sq.</i> | <i>F</i> |
|--|-----------|--------------|--------------------------|
| Redundancy | 1 | 1.00 | 0.37 at 1 + 10 <i>df</i> |
| Task complexity | 2 | 0.75 | 0.65 at 2 + 10 <i>df</i> |
| Groups | 5 | 1.40 | 1.22 at 5 + 10 <i>df</i> |
| Redundancy \times complexity | 2 | 3.59 | 3.12 at 2 + 10 <i>df</i> |
| Redundancy \times groups | 5 | 1.07 | 0.93 at 5 + 10 <i>df</i> |
| Complexity \times groups | 10 | 1.05 | 0.91 at 5 + 10 <i>df</i> |
| Redundancy \times complexity \times groups (error term) | 10 | 1.15 | |

TABLE 3
MEAN REACTION TIME IN SECONDS FOR THE SIX EXPERIMENTAL CONDITIONS

| | 2 elements | 3 elements | 4 elements |
|--------------|------------|------------|------------|
| Redundant | .568 | .761 | .650 |
| Nonredundant | .432 | .528 | .585 |

TABLE 4
MEAN POSITIVE ERROR SCORES FOR THE SIX EXPERIMENTAL CONDITIONS

| | 2 elements | 3 elements | 4 elements |
|--------------|------------|------------|------------|
| Redundant | 2.33 | 1.17 | 1.00 |
| Nonredundant | 0.83 | 1.00 | 1.66 |

D. DISCUSSION

The results clearly support the two hypotheses based on the RT data, namely that RT is longer with redundant than with nonredundant organization, and RT increases as task complexity increases. These data are also in agreement with those of Baxter (2) and Saltzman and Garner (5).

The experiment provided insufficient data to test the two hypotheses regarding negative errors:

1. Fewer negative errors occur in the redundant condition.
2. The number of negative errors increases as stimulus complexity increases, and this increase is greater for the nonredundant condition.

It is felt that this failure to obtain data on negative errors indicates that in all cases the level of task difficulty was too low. This weakness in the design can be corrected by either increasing the number or the complexity of the critical stimuli or reducing the exposure duration of the stimuli.

The failure to obtain negative errors when combined with the occurrence of a relatively large number of positive errors leads to the conclusion that at least under the experimental conditions employed here, the probability of positive errors was considerably greater than that of negative errors. The

reason for this difference is a subject for further investigation. It seems reasonable to hypothesize that although the instructions regarding positive and negative errors appear to be equal, the emphasis on speed of response might have been responsible for the excess of positive errors.

Although sufficient positive errors were obtained to allow data analysis, none of the effects differed significantly from a chance expectation (minimum level for rejection of the null hypothesis was five per cent). However, an interesting, if somewhat confusing, trend can be seen in Table 4. As predicted there are more positive errors for the redundant condition at the low and intermediate levels of difficulty; however, at the highest level of difficulty employed, the nonredundant condition shows more errors than the redundant (i.e., the trend is in the direction *opposite* that predicted). If future investigations employing a greater range of task difficulty should support this trend as well as validating the hypotheses concerning negative errors, it would provide justification for designing systems for redundant operation which have the advantages not only of fewer negative errors but also, if the task is sufficiently complex, fewer positive errors. Both of these advantages would, of course, be purchased at a cost of longer RT's, however, RT-error tradeoffs could be established for varying levels of task difficulty (as measured by stimulus complexity, number of critical stimuli, stimulus duration, et cetera).

E. SUMMARY

The study investigated speed and reliability of performance by pairs of human operators as a function of operator redundancy and task difficulty. Six pairs of subjects responded to three sets of four critical stimuli composed of either two, three, or four lights in a 4×3 matrix of lights; the 12 "critical stimuli" were randomly placed in a total set of 48 stimuli. Pairs of subjects were arranged for redundant or nonredundant operation. In the redundant operation, both subjects responded to all four "critical stimuli" with the fastest response of the two being recorded. In the nonredundant situation, each subject responded to two of the four "critical stimuli." All subjects served under all experimental conditions with the order of presentation of conditions counterbalanced between groups.

Results substantiated the hypothesis that reaction time would decrease with decreasing stimulus complexity and that the nonredundant situation would produce a lower reaction time than the redundant condition. Analysis of number and types of errors showed no significant effects.

REFERENCES

1. ADAMS, J. A. Vigilance in the detection of low-intensity visual stimuli. *J. Exper. Psychol.*, 1956, **52**, 204-208.
2. BAXTER, B. A study of reaction time using factorial design. *J. Exper. Psychol.*, 1942, **31**, 430-437.
3. BESGUM, B. D., & KLEIN, C. I. A survey and analysis of vigilance research. Report No. 8, Human Resources Research Office, Washington, D.C., 1961.
4. GOTTSACKER, R.^o Reaction time: The time to initiate a response. M. H. Aero Doc. U ED 6102, Minneapolis-Honeywell Regulator Co., Minneapolis, 1958. Pp. 13.
5. SALTZMAN, I. J., & GARNER, W. R. Reaction time as a measure of span of attention. *J. of Psychol.*, 1948, **25**, 227-241.
6. SCHAFER, T. H. Detection of a signal by several observers. NE 121301, USNEL Report No. 101, Jan., 1949.
7. WILLIAMS, S. B. The Search Factor in Detecting Weak Radar Targets. Baltimore, Md.: Johns Hopkins University, Psychology Laboratory, 1947 (July).

Sylvania Electronic Systems

A Division of Sylvania Electric Products, Inc.

Applied Research Laboratory

40 Sylvan Road

Waltham 54, Massachusetts



RESPONSE SET IN OBJECTIVE ACHIEVEMENT TESTS*¹

New York University

ALICE GUSTAV²

A. INTRODUCTION

A considerable body of literature exists concerning the phenomenon of response set in self-report scales. Cronbach (4, p. 372) who now calls it response style, defines it as "... a habit or momentary set which causes the subject to earn a different score from the one he would earn if the same items were presented in a different form." Response set was originally described by Lorge (7) in connection with the Strong Vocational Interest Blank. Subsequently its presence, causes, and effects have been extensively explored in personality, interest, and attitude scales. But surprisingly little work has been reported in the case of objective tests of intellectual material, despite the astronomical numbers of such tests administered annually, and despite the crucial decisions frequently made on the basis of such tests. Yet, if these tests were to be found susceptible to response set, their validity would have to be reconsidered.

Items on objective tests of knowledge and ability are, most usually, in true-false and/or multiple-choice format. Cronbach (4, p. 372) has warned that response set is an important factor in true-false tests:

In true-false tests particularly, some people have the habit of saying "true" when in doubt, while others are characteristically suspicious and respond "false" when in doubt. If the tester has included a large proportion of true statements in his test, the acquiescent student will earn a high score even if his knowledge is limited.

In connection with the multiple-choice format there seems to be a complacent acceptance of its impregnability. A search of the literature unmasked only one study on response set in multiple-choice tests of knowledge. Cronbach (3, p. 10), using the Henmon-Nelson Test of Mental Ability, Form A, and a modified version of the Ohio State University Psychological Examination,

* Received in the Editorial Office on June 14, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ This project was partially supported by a grant from the New York University Arts and Science Research Fund.

² Grateful acknowledgment is made to Miss Carolyn Feinberg for the painstaking tabulation of responses from the original test material.

came to the conclusion that there was so little response set in multiple-choice tests that it would be desirable to "... encourage their continued widespread use." And, in his latest text Cronbach (4, p. 372) now says:

For measuring ability, multiple-choice or best-answer tests are distinctly preferable to tests having fixed response categories such as true *vs.* false or agree *vs.* disagree. The best-answer test is not only virtually immune to response biases other than tendency to gamble but is especially well adapted to testing of comprehension.

The present author's compulsive attention to construction of classroom tests and item analyses of students' response to those tests led to the suspicion that response set does manifest itself in multiple-choice as well as in true-false formats. Several facts bolstered this feeling. (a) In constructing the items, either by making them up herself or using items from manuals provided by textbook authors, there was a strong tendency *not* to wind up with an equal number of A's, B's, C's, and D's as the correct options. Alternatives within items frequently had to be rearranged so that each option would appear as the correct one an approximately equal number of times in a given test. This was supported by Berg and Rapaport (1) who found that when students were asked to prepare multiple-choice items, of 1000 consecutive questions they constructed:

... the correct answers were located as follows: 1, 54; option 2, 237; option 3, 571; option 4, 138. For significance at the one per cent level a chi-square value of only 11.3 would be required for these data. However, the actual chi-square value for this distribution of answer position reaches the fantastic level of 616.7, a convincing demonstration of bias in a free situation (1, p. 478).

(b) In the item analyses done on students' responses, there began to be evidence that errors were not being distributed in a random fashion. This also occurred in the Berg and Rapaport study (1) in which definite evidence of response set was found even with a multiple-choice format on a questionnaire which had no actual questions but which required the subjects to imagine the correct answers when the various options were presented to them.

These clues prompted a decision to investigate systematically the possibility of response set in objective classroom tests, i.e., an area of testing extensively used on the American educational and vocational scenes.

B. PROCEDURE

Twelve tests which had been administered as final examinations to undergraduate psychology classes were used. Ten were prepared by this author,

with some items constructed by herself and some taken from manuals accompanying assigned textbooks. The remaining two tests were of the same nature but were prepared by, and administered to the classes of, another member of the department. The subject matter of the courses included introductory psychology, emotion and motivation, adolescent psychology, and abnormal psychology. All twelve tests consisted of both true-false and multiple-choice items with four alternatives (options A, B, C, D). Instructions required that the students mark the former items either "T" or "F" and in the case of the latter items to write the letter of the one best alternative. The tests were scored simply for the number of correct answers. No corrections for chance were made, as it has long been the policy of this author not to use corrections for chance which are, in essence, penalties for boldness (5, 9). Since these were final examinations, the students were fully familiar with this procedure and had learned to attempt every item without fear of extra penalty for being wrong.

There were 520 students (266 men and 254 women). Results were very similar for men and women, so they have been combined in presentation of the statistical material below.

Theoretically, the wrong answers for each student should contain the possible options in the same proportions as the options actually occur in the total test key of correct answers. For example, if a given test had 10 per cent each of true and false items, and 20 per cent each of A, B, C, and D items, then we would expect that a student's wrong answers should also include 10 per cent each of T and F responses and 20 per cent each of A, B, C, and D responses. Any consistent deviation from this pattern would be evidence of response set. Therefore, the first step was to calculate the actual number of times each option appeared in the test keys and convert these into percentages of the total test. Next, the wrong answers for each student were tabulated in terms of his response, i.e., if the item was wrong, did he mark it T, F, A, B, C, or D? The number of wrong answers for each student was totalled, and percentages of that figure were computed for each of his incorrect response options. Finally, the per cent for each option for each student was compared with the test key per cents to ascertain whether he deviated consistently on any. These were tabulated under three headings: "more," "same," and "less." The category "same" consisted of a range of three around the per cent on the test key for a given option. In many cases this meant approximately the middle third of the distribution, as in cases, for example, where there were only 10 per cent T or F items on a test. The categories of "more" and "less" contained per cents which were

above or below this range of three. The number of students in each category was combined for all 12 tests, and levels of significance were computed.

C. RESULTS

Table 1 summarizes the results. It is read as follows: 321 (62 per cent) of the students marked their wrong answers "T" a larger per cent of the time than "T" actually appeared on the test keys, 49 (10 per cent) of the students marked their wrong answers "T" in the same proportion as the test keys, and 150 (28 per cent) marked their wrong answers "T" in a smaller per cent of time than it actually occurred on the test keys. It can be seen that there is a strong trend for students to give larger percentages of T, A, and B, as incorrect responses and to give smaller percentages of F, C, and D as incorrect responses.

TABLE 1
DISTRIBUTION OF STUDENTS' WRONG RESPONSES COMPARED WITH DISTRIBUTION
OF CORRECT RESPONSES ON TEST KEYS

| Answer option | More than test key N % | Same as test key N % | Less than test key N % | χ^2 Hypothesis of normal distrib. in total group | χ^2 Hypothesis of equality of more & less groups |
|------------------|------------------------------|----------------------------|------------------------------|--|---|
| T | 321 62 | 49 10 | 150 28 | 995.69* | 62.08* |
| F | 59 11 | 43 9 | 416 80 | 1607.56* | 268.32* |
| A | 241 46 | 79 15 | 200 39 | 676.50* | 3.82** |
| B | 252 48 | 91 18 | 177 34 | 543.22* | 13.12* |
| C | 179 34 | 85 17 | 256 49 | 673.23* | 13.62* |
| D | 191 37 | 74 14 | 254 49 | 709.84* | 8.92* |

* Significance level $< .01$.

** Significance level $> .05$.

As already indicated, one would assume that the frequency of wrong answers using any given option would be the same as the frequency of that option on the test key. At least, allowing for casual errors, one would expect something approaching a normal distribution, with the bulk of wrong responses in the center, and a tapering off in the "more" and "less" ends of the distribution. To test for this, chi squares were calculated for each option and, as can be seen from Table 1, the extraordinarily high chi squares indicate that the hypothesis of a normal distribution must be rejected and that the deviations from such a theoretical normal distribution are highly significant.

Granted that there is a significant deviation from an assumed normal frequency and that there are, essentially, bimodal distributions for each option, is there any definite expectation that the "more" or "less" category

will prevail? Theoretically, we might expect that of those students who deviated from the "same" pattern, half would be in the "more" and half in the "less" direction. Chi squares were again resorted to as a test and, as can be seen from Table 1, with the exception of alternative "A," the hypothesis of such an assumed equal distribution must be rejected, as the deviations therefrom are highly significant. Even option "A" almost reaches the 5 per cent level of significance. A chi square of 3.841 would have been significant at the 5 per cent level, and the obtained chi square is 3.82.

D. DISCUSSION

The strong response set evident in true-false items supports Cronbach's statement quoted earlier, in which he describes some people as more acquiescent and therefore responding more often with "T," or more suspicious and hence responding more frequently with "F." Couch and Keniston (2) have postulated the interesting hypothesis that:

Yeasayers were shown to be individuals with weak ego controls, who accept impulses without reservation and who 'agree' and easily respond to stimuli exerted on them. The naysayer inhibits and suppresses his impulses, in many ways rejecting all emotional stimuli impinging on him (2, p. 173).

Perhaps the great preponderance of "yeasayers" among the students of this study constitutes further evidence of the effects of our culture on the value orientations of college students as summarized by Jacob (6). He has pointed out that we have raised a generation of young people who are incredibly self-centered and contented. It may well be that in tests, as in other areas of life, they do not trouble themselves to question or to disagree.

The most interesting finding in this study is the definite presence of response set in the multiple-choice portion of the tests.

Of course, it is entirely possible that the alternatives in multiple-choice questions on a test might be constructed in such a fashion that a given option is always a more difficult one. This might throw a greater number of wrong responses to the other options. In that case, it would be a problem of response set on the part of the test constructor. The present author has always been aware of this eventuality and has deliberately counteracted it by rearranging options in a random order and by providing for an approximately equal number of correct options in a given test. This eliminates possible author bias. In addition, the results here are based on 12 different tests, including two from another teacher.

The amount of response set found in the present data is, if anything, a minimum. In his study Cronbach (3) said that the:

. . . papers of high-scoring pupils (those having a score above 60 out of 90 items correct) were discarded. This was done to *increase* the likelihood of finding a response set since response sets have no opportunity to show themselves when the pupil gets most items correct (3, pp. 8-9).

In this study no papers were discarded. Since having the papers of high-scoring students would decrease the presence of response set, it highlights even more dramatically the results described in the preceding section.

As for the preponderance of A and B over C and D in the multiple-choice items, Cronbach (3) originally suggested

. . . that some students may persistently tend to select choices early in the group of five. . . . The psychological basis for the hypothesis is the possibility that some students read every alternative and discriminate carefully, where some merely read through the item to find a plausible answer, mark it, and go on to the next item (3, p. 8).

Later in the same study he rejected this when he found practically no evidence of response set. The present investigation would seem to indicate that his original hypothesis may have merit.

As this paper was being prepared for publication, an article by Sax and Carr (8) appeared, in which they reported that students obtain higher scores on the Henmon-Nelson Tests of Mental Ability when the material is presented in a spiral-omnibus, rather than in a subtest, form. They concluded that this was to be considered a response set.

In view of all of the foregoing it might be wise to initiate a more penetrating evaluation of the validity of the multiple-choice format. The fact that it has become so well-entrenched in our culture makes it accepted without question. Such blind conformity ought not be encouraged by psychologists.

E. SUMMARY

Wrong answers on objective classroom examinations of 520 students in undergraduate psychology classes were tabulated and the per cent of wrong answer options compared with the actual frequency of options on the test keys. The following statistically significant deviations were found: Students tended to make greater use of T, A, and B as wrong answer options and to make lesser use of F, C, and D alternatives as wrong answer options. The conclusion is drawn that response set is a factor which must be taken into

account in connection with multiple-choice as well as true-false tests, and in tests of knowledge and ability as well as in self-report tests.

REFERENCES

1. BERG, I. A., & RAPAPORT, G. M. Response bias in an unstructured questionnaire. *J. of Psychol.*, 1954, **38**, 475-481.
2. COUCH, A., & KENISTON, K. Yeasayers and naysayers: Agreeing response set as a personality variable. *J. Abn. & Soc. Psychol.*, 1960, **60**, 151-174.
3. CRONBACH, L. J. Further evidence on response sets and test design. *Educ. Psychol. Meas.*, 1950, **10**, 3-31.
4. ———. *Essentials of Psychological Testing* (2nd ed.) New York: Harper & Bros., 1960.
5. DOPPELT, J. E. The correction for guessing. *Test Service Bulletin* (The Psychological Corporation), 1954, **46**, 1-4.
6. JACOB, P. E. *Changing Values in College*. New York: Harper & Bros., 1957.
7. LORGE, I. Gen-like: Halo or reality. *Psychol. Bull.*, 1937, **34**, 545-546.
8. SAX, G., & CARR, A. An investigation of response sets on altered parallel forms. *Educ. Psychol. Meas.*, 1962, **22**, 371-376.
9. SHERRIFFS, A. C., & BOOMER, D. S. Who is penalized by the penalty for guessing? *J. Educ. Psychol.*, 1954, **45**, 81-90.

Department of Psychology
New York University
Washington Square
New York 3, New York



PROBABILITY PREFERENCES AND EXPECTED VALUES*¹

Department of Psychology, The Ohio State University

ALVIN SCODEL

A. INTRODUCTION

In a series of experiments some years ago Edwards (1, 2, 3, 4) demonstrated that subjects display stable preferences for high or low probability bets when expected dollar values are equal, variance preferences are a minor factor in determining betting behavior as compared to probability preferences, and probability preferences are highly instrumental in determining chances even when expected values vary. The stability of these probability preferences and the relative unimportance of variance preferences are disputed by McGlothlin (6) from his analysis of horse-race betting. McGlothlin's most general finding was that racetrack bettors would accept low expected dollar values when low probability-high prizes were involved and, conversely, would demand higher expected values in cases of high probability-low prize combinations. This finding is consistent with the results of Preston and Baratta (7). In an auction situation involving pairs and sets of four players they found that prizes with low probabilities were paid for too generously and prizes with high probabilities were taken as bargains. When the number of players increased from two to four, there was an increase in the over and under valuation of prizes at extreme probability values.

Scodel, Ratoosh and Minas (8) investigated preferences for bets in groups of college students and military subjects. The payoff matrix contained nine alternative bets, three each of positive, negative, and zero expected value. While there were pronounced differences between the groups in their preferred bets, expected value appeared to be a negligible factor within the limits of this investigation in determining the preferences of either group.

Needless to say, the results which have been cited are exceedingly ambiguous concerning the relative importance of probability preferences and expected values in determining bets. The present study, part of a larger attempt to correlate betting preferences with personality variables, utilized

* Received in the Editorial Office on June 17, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ This research was supported in whole by the United States Air Force under Contract No. AF49(638)-317 monitored by the AF Office of Scientific Research of the Air Research and Development Command.

a payoff matrix in which preferred bets, as determined by a previous study, were given negative expected value and least preferred bets were given positive expected value.

B. METHOD

Sixty male Ss from the introductory psychology class at Ohio State University participated in this experiment. In the risk-taking situation each S was given \$10.00 and instructed as follows:

You are going to play dice. You can't lose any money, but you have a very good chance to make some. You are going to throw the dice 50 times. Each time you will bet \$.20. Before you bet on each throw, tell me which of the seven possible bets you have selected.

The payoff matrix, Table 1, was shown to S and left in view during the remainder of the session.

TABLE 1
THE PAYOFF MATRIX

| Bet: | A | B | C | D | E | F | G |
|-----------|-----------------|---------------------|--------------|--------|--------|--------|---------|
| Wins on: | 5 6 7 8 9 | 2 4 6 8 10 12 | 8 9 10 | 5 | 7 | 9 | 3 |
| .20 pays: | .05 | .10 | .30 | .60 | 1.00 | 1.75 | 3.90 |
| Chances: | 2 in 3 | 1 in 2 | 1 in 3 | 1 in 4 | 1 in 6 | 1 in 9 | 1 in 18 |

Objective probabilities were carefully explained, and numerous examples were given to clarify the odds. It was emphasized that the payoff did not include the amount bet. S was then told:

At the end of the 50 throws, if you have \$10.00 or less, you keep 10% of that amount. For example, if you end up with \$8.00, you keep \$.80. If you end with more than \$10.00, you keep \$1.00 plus half of everything over \$10.00. For example, if you end up with \$13.00, you would keep \$1.00 plus one-half of \$3.00, or \$2.50.

The situation was not rigged in any way. Outcomes were determined entirely by chance and wins ranged from \$.30 to \$6.20.

In a previous study (5) a matrix identical in some respects to the present one had been employed. However, in the earlier study there were four possible amounts that could be bet (\$.10, .20, .30, or .40), and each bet had zero expected value. Under these conditions the preferred bets were B, C, and A in that order (objective probabilities of 1/2, 1/3 and 2/3). The least preferred bets were G and F (objective probabilities of 1/18 and 1/9). Ss were not given information in the earlier study or in this one about the ex-

pected values of the various bets although, as indicated, the objective probabilities were carefully explained. *Ss* were, of course, free in both experiments to determine the expected values of bets for themselves if they wished, but it was felt that any information from the experimenter about the positive and negative expected values of bets would seriously bias the choices of bets.

The modal bet of the earlier study (the B bet with 0.5 objective probability) paid off only \$.10 for \$.20 bet in this study, i.e., it had a negative expected value of \$.05. The other two preferred bets, C and A, were each given negative expected values of \$.03 $\frac{1}{3}$. The two least preferred bets, G and F, were given positive expected values of \$.02 $\frac{7}{9}$ and \$.01 $\frac{2}{3}$ respectively. The D and E bets had zero expected value.

Ss were not instructed to maximize monetary return. Presumably such instructions could focus more attention on the expected values of bets. The intent of this study was simply to see if preferred bets remained preferred despite a manifestly less than fair payoff (particularly in the case of the B bet) and, conversely, if ordinarily avoided bets would be taken more often if they were given positive expected values. Obviously, expected values could be made so extreme as to necessarily have a pronounced influence on betting preferences despite the absence of any instructions to maximize return.

C. RESULTS

*The frequencies of bets are given in Table 2.

The most frequent bets, C and B, are both bets with negative expected value. Actually, the B bet was the modal bet for 19 *Ss* as compared to 17 *Ss* for whom the C bet was modal. G and F, the two bets with positive expected value and with the highest payoffs, ranked fifth and seventh in fre-

TABLE 2
* FREQUENCIES OF BETS ($N = 60$)

| | A | B | C | D | E | F | G |
|--------|------|-------|-------|------|------|------|------|
| Total | 262 | 614 | 666 | 414 | 458 | 254 | 332 |
| Mean | 4.37 | 10.23 | 11.10 | 6.90 | 7.63 | 4.23 | 5.53 |
| Median | 2 | 8 | 9 | 5 | 5 | 2 | 2 |

quency among the seven possible alternatives. The preferences for bets cannot be compared directly with those of the Liverant-Scodel study since that study provided only 30 trials for each *S* and also allowed for variability in amounts that could be bet, but, as in that study, B and C remained the most preferred bets whereas F and G were generally avoided. It seems clear that, within the range of expected values used in this study, such values are

negligible in determining betting preferences. In general, the preferences provide striking evidence for the conservatism in betting of college Ss. The only contrary evidence is the avoidance of the A bet which had an objective probability of two in three, but the absence of challenge in a bet with such a high objective probability would seem to be a plausible explanation for the avoidance of this bet.

D. DISCUSSION

While the results indicate support for Edwards' finding that probability preferences play a major role in influencing choices even when expected values vary, there is still the problem of reconciling these results with those of the McGlothlin and Preston and Baratta studies. Both of these studies found that Ss accept low expected values for low probability—high payoff outcomes, but demand higher expected values for outcomes with high probability and low reward. The high probability-low reward outcome is undervalued, and the low probability-high reward outcome is overvalued. In the present study Ss retain a preference for high probability-low payoff bets even when these bets have negative expected value and avoid low probability-high payoff bets though they have positive expected value. Differences in Ss hardly explain the discrepancies in results since the McGlothlin study analyzed horse-race bets, and Preston and Baratta employed Ss in an academic setting drawn from students and faculty. What was distinctive about both these studies, however, was that Ss made bets in a group situation. In the Preston and Baratta experiment competitive bids were made for prizes, and McGlothlin's data were derived from one day's betting behavior at a track.

It is not the purpose of this report to become involved in the arcana of mathematical decision-making theory, but certainly it is worth considering that an important variable to be incorporated into decision-making theory is the effect of group facilitation on choices. One hypothesis which emerges from the cursory attempt to reconcile discrepant results in this area is that the presence of other people engaged in the same kind of decision-making behavior—whether one is supposed to be competing overtly with them or not—produces less conservatism in choices. To test this hypothesis, seven pairs of Ss, all males from the introductory psychology course at Ohio State, were run through the same matrix that was employed in this study. Instructions were precisely the same, the only variation being that Ss alternated in making their bets. No communication between members of a pair was allowed, but each heard the other's selection of bets as they were announced to the experimenter and witnessed the outcomes of the bets. For these 14 Ss the means of the

seven alternative bets were as follows: A, 0.57; B, 6.64; C, 10.64; D, 9.36; E, 8.43; F, 5.00; G, 9.29. If these results are compared with the frequencies presented in Table 2, it can be seen that in the two-person situation the three most conservative bets, A, B, and C, decrease in frequency while all others increase. (A Mann-Whitney comparing the sums of the A, B, and C bets for *Ss* in the two groups results in a z of 1.74, $p = .04$). Obviously, this change cannot be attributed to a disposition to shift from bets with negative expected value to bets with positive expected value since the D and E bets, which have zero expected value, also increase in frequency; moreover, the C bet which has negative expected value is the modal bet for more *Ss* than any other bet (5). The most likely explanation is that the presence of another person produced a competitive set (the procedure of alternating bets may well have intensified this set) even though the instructions said nothing about attempting to win more money than the other *S*. A "rational" strategy designed to maximize the difference in payoff to oneself and the other person would dictate a concentration on bets with positive expected value (F and G), but *Ss*, more or less impervious to expected values, simply shift all of their betting in the direction of higher payoffs. That is, the need to do as well or better than the other person means that less utility is placed in winning per se, and greater utility is now attached to payoffs. The competitive set makes the number of favorable outcomes less important than the need to accrue larger payoffs since accumulated money is the criterion the person uses to assess his performance vis-à-vis the other person. The actual betting behavior seems to be a compromise between the conservatism of most *Ss*, which emphasizes favorable outcomes rather than payoffs, and the fear that the other person will accumulate more money unless one now plays more "long shots." If these speculations are correct, they would explain why, in a group betting situation, high probability-low payoff combinations are undervalued and low probability-high payoff combinations overvalued, independent of their expected values.

E. SUMMARY

Sixty male undergraduate *Ss* participated in a gambling situation in which each *S* was required to bet on the outcome of the toss of a pair of dice 50 times. On each trial *S* selected a bet from seven alternative outcomes with known objective probabilities but different expected values. Bets known to be preferred from a previous study were given negative expected value, and least preferred bets were given positive expected value. *Ss* were given money with which to bet and were told that they could keep the money they won.

Expected values apparently had little, if any, effect in determining selections of bets. Conservative bets were preferred even though they had appreciable negative expected values and "long shot" bets were generally avoided although they had positive expected values.

Subsequent to the collection of these data, the hypothesis was tested that two people making bets, under the same conditions of instructions and pay-offs and with knowledge of each other's bets and outcomes, would make more low probability-high payoff bets than Ss run individually. This hypothesis was confirmed. Speculations were offered to reconcile differences in the results of several decision-making studies.

REFERENCES

1. EDWARDS, W. Probability preferences in gambling. *Amer. J. Psychol.*, 1953, **66**, 349-364.
2. ———. Probability preferences among bets with differing expected values. *Amer. J. Psychol.*, 1954, **67**, 56-67.
3. ———. The reliability of probability preferences. *Amer. J. Psychol.*, 1954, **67**, 68-95.
4. ———. Variance preferences in gambling. *Amer. J. Psychol.*, 1954, **67**, 441-452.
5. LIVERANT, S., & SCODEL, A. Internal and external control as determinants of decision making under conditions of risk. *Psychol. Rep.*, 1960, **7**, 59-67.
6. MCGLOTHLIN, W. H. Stability of choices among uncertain alternatives. *Amer. J. Psychol.*, 1956, **69**, 604-615.
7. PRESTON, M. G., & BARATTA, P. An experimental study of the auction-value of an uncertain outcome. *Amer. J. Psychol.*, 1948, **61**, 183-193.
8. SCODEL, A., RATOOSH, P., & MINAS, J. S. Some personality correlates of decision making under conditions of risk. *Behav. Sci.*, 1959, **4**, 19-28.

Department of Psychology
Ohio State University
1945 North High Street
Columbus 10, Ohio

A STUDY OF THE POSSIBLE EFFECT OF DIFFERENTIAL LIGHT ADAPTATION ON MEASURES OF BRIGHTNESS ENHANCEMENT*¹

Department of Psychology, Michigan State University

CHARLES M. BOURASSA AND S. HOWARD BARTLEY

A. INTRODUCTION

The measurement of brightness enhancement involves the simultaneous presentation of two target areas, one steadily illuminated and the other intermittently illuminated. The viewing of these targets involves light adaptation. Since, per unit time, the intermittent target delivers less photic flux to the eye than the steady target does, it may produce less light adaptation (6). Were this to be the case, light adaptation would be a factor in determining the calculated amount of brightness enhancement (6).

Brightness enhancement has been accounted for, in principle, by the particular timing given to the intermittent photic input. The activity of the central part of the optic pathway involved in producing brightness enhancement has been depicted by Bartley's alternation of response theory (1, 2, 3, 4, 5, 7, 8, 9), a theory that not only accounts for brightness enhancement, but reduction in visual acuity and several other visual end results under intermittent illumination.

It was thought advisable to test for a possible production of differential amounts of light adaptation in the two targets under conditions simulating those used in typical brightness enhancement investigations. The present study was such a test and was not a precise study of light adaptation as such. In other words, the problem lay in determining whether the mere difference in flux per unit time in the standard and comparison targets biased the results in such a way as to produce any considerable part of the effect taken to be dependent upon intermittency.

B. PROCEDURE AND APPARATUS

Accordingly the following conditions were set up. (For a picture of these, see Figure 1.) I. Two similar targets, both 13 seconds in duration were

* Received in the Editorial Office on June 17, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ The present paper is a part of the work being done under Research Grant NSF-G-19485 from the National Science Foundation.

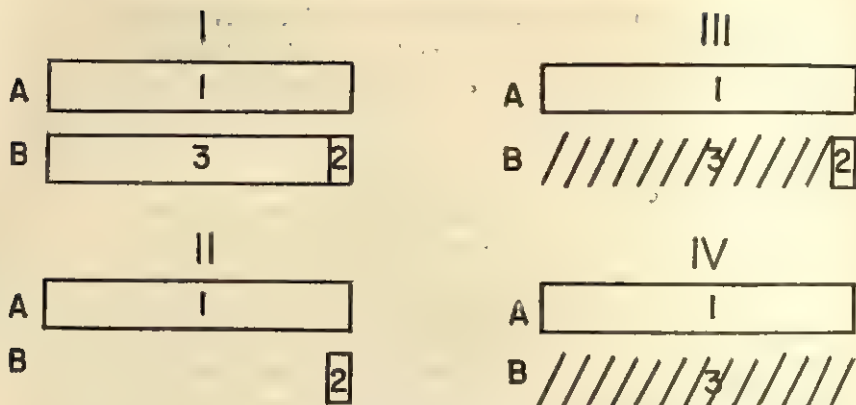


FIGURE 1

SCHEMA SHOWING THE VARIOUS TEMPORAL FEATURES OF IMPINGEMENT ("STIMULATION")

A and B are the two adjacent targets. Arabic figures indicate photic sources used as described in the text. The slash lines in IIIB and IVB indicate intermittent impingements. The horizontal axis of the bars represents time.

presented. These were to be matched in brightness by the observers as a control condition. II. Two targets, one 13 seconds long, the other 1 second long, concurrent with the last second of the first. The intensity of the second was adjusted so that the brightnesses of the two targets matched. III. Two targets, one 13 seconds long with steady illumination, the other 12 seconds with intermittent illumination, and the final one second with steady illumination. Here again the two targets to be matched in brightness were steady, but the short target was preceded by intermittent stimulation. With this arrangement, it was not brightness enhancement that was measured but rather the possible adapting effects of intermittent stimulation. IV. Two 13-second targets, one steady, the other one intermittent. Here brightness enhancement was measured. In all cases 20 seconds elapsed between each stimulus presentation.

The apparatus was as follows: Three photic sources were used for the two targets, A and B. These were housed so that the only emission was through opal glass windows. The sources, 1 and 2, were in the observer's line of regard and 2 was seen through a half-silvered mirror. The third source, equipped with an interchangeable episcotister was at right angles to the other two, and seen via the half-silvered mirror used for source 2. Thus source 1 provided target A and either source 2 or 3 provided target B. Target A was always steady, a standard for matching target B. Target B could be made steady or intermittent as needed for the conditions already specified.

In one part of the study, the level of the standard target A was set at 25 c/ft² and in the other it was set at 50 c/ft². Intensity was established through control of lamp current. This was thought adequate since the needed total intensity range was not great. The durations and temporal separations of target exposure were controlled by cam-activated micro-switches.

To eliminate as much of the stray illumination as possible and to control target size, a reduction screen was placed between the equipment just described and the observer 30 cm away. The targets were 2.3 cm squares with an intertarget separation of 3.3 cm. The visual angles subtended by the targets were 4.39°, with the interspace subtense, 6.3°.

Further to simulate brightness enhancement experimentation, three pulse-to-cycle fractions (PCFs) were used, namely 0.7, 0.5, and 0.3. The rate of intermittency was measured by a Weston tachometer. The intermittency rate was maintained at 10 cycles per second throughout the work. Under most conditions this rate is likely to produce maximal brightness enhancement. Twenty brightness matches were made under each condition, including the three PCF conditions in III. Since the one-second presentation of the test stimulus precludes the possibility of the observer actively adjusting its intensity while it is on, intensity adjustments leading to a brightness match were made during the null period between stimulus presentations. There were two observers, one (C) experienced; and the other (N), an undergraduate without experience and without knowledge of the purpose of the investigation.

The matches were difficult and several hours of preliminary practice were required.

C. RESULTS AND DISCUSSION

Although brightness enhancement is the brightness by which an intermittent target exceeds that of a steady one of the same luminosity, the measures which generally represent it are not given in terms of brightness units but by a *brightness index*. This is simply the ratio of the increased intensity of the steady target to the intensity of the intermittent one. This obviously tells nothing about actual brightness levels of either target but does indicate the relative effectiveness of intermittent illumination as compared to steady illumination.

In Figure 2, the open bars in both C and N indicate the mean brightness indexes for condition IV, the usual circumstances for producing brightness enhancement. The left hand portion for each observer is for the higher intensity of the standard. It will be noted that as PCF is reduced, brightness enhancement increases. This is true for both observers C and N. The closed

PCFs .5 and .3 progressively less so. This would suggest that the effect of the pretest interval is not solely a possible photochemical adaptation period but a period during which the temporal features of the photic input influence the neural outcome, the influence being opposite to that expected from light adaptation.

However, at the higher intensity, the results are less internally consistent. For observer C, the inconsistency is marked. For observer N, the results are simply opposite to those at the lower intensity level. This may suggest that for this observer, the greater intensity was beginning to be enough for light adaptation to have more influence.

Owing to certain variables inherent in observing under such conditions, inconsistencies of this order are not surprising. The conclusions we can draw from the work are the following:

1. Whether the findings in conditions II and II are due to light adaptation or to the temporal features of photic input which influence neural activity as described by the alternation of response theory (1, 2, 4, 5, 9), the effects are slight in comparison to the magnitude of brightness enhancement itself, as shown in condition IV.

2. Thus what is called brightness enhancement here and in previous studies is a genuine phenomenon not primarily to be accounted for by light adaptation. Further than this we cannot and need not go. The intent of the study was simply to check on the possible existence and relative magnitudes of some heretofore untested factor such as light adaptation as a possible basis for brightness enhancement. The major basis for brightness enhancement seems to lie outside of adaptation.

3. The study provided no evidence to negate the supposition that brightness enhancement is a phenomenon of impingement timing in relation to the cyclic characteristics of the central portion of the optic pathway, a supposition based on a number of neurophysiological findings and subsequent confirmations of theory projected from the findings.

REFERENCES

1. BARTLEY, S. H. A central mechanism in brightness enhancement. *Proc. Soc. Exper. Biol. & Med.*, 1938, **38**, 535-536.
2. ———. Some effects of intermittent photic stimulation. *J. Exper. Psychol.*, 1939, **25**, 462-480.
3. ———. Some factors in brightness discrimination. *Psychol. Rev.*, 1939, **46**, 337-358.
4. ———. The relation between cortical response to visual stimulation and changes in the alpha rhythm. *J. Exper. Psychol.*, 1940, **27**, 624-639.
5. ———. Visual sensation and its dependence upon the neurophysiology of the optic pathway. *Biol. Symp.*, 1942, **7**, 87-106.

6. ———. Light adaptation and brightness enhancement. *Percept. & Motor Skills*, 1957, 7, 85-92.
7. ———. Some facts and concepts regarding the neurophysiology of the optic pathway. *A.M.A. Arch. Ophthalm.*, Part II, 1958, 60, 775-791.
8. ———. Central mechanisms in vision. In Field, J., Magoun, H. W., & Hall, V. G. (Eds.), *Handbook of Physiology, Sec. 1, Neurophysiology, Vol. I*. Washington, D. C.: Amer. Physiol. Soc., 1959. Chapter 30, pp. 713-741.
9. ———. A clarification of some of the procedures and concepts involved in dealing with the optic pathway. In Jung, R., & Kornhuber, H. (Eds.), *The Visual System: Neurophysiology and Psychophysics*. Heidelberg: Springer-Verlag, 1961.

Laboratory of Neurophysiology
Good Samaritan Hospital
Portland 10, Oregon

Department of Psychology
Michigan State University
East Lansing, Michigan

ANXIETY, EXTRAVERSION, AND PERSONALITY IDIOSYNCRASY IN DELINQUENCY*

*Green Hill School, Chehalis, Washington, and Fort Worden School, Port Townsend,
Washington*¹

GEORGE R. PIERSON AND ROBERT F. KELLY

A. INTRODUCTION

Recent application of factored personality tests, and specifically the High School Personality Questionnaire (hereafter called the HSPQ), to the understanding and diagnosis of delinquency has led to the development of both a profile and a formula, DELST to distinguish potential for delinquent behavior (4, 7, 8). DELST is a single score of delinquent potential derived from a combination of HSPQ factor scores. However, more than this is needed to analyze the delinquent and indeed more can be obtained from the HSPQ by further use of factor scores. This paper is a report of such use.

It has been observed repeatedly, both in clinical and research work with delinquents, that extreme variation on certain dimensions is characteristic of delinquency. Sir Cyril Burt (2) of England, an early pioneer in the application of factor theory to the problems of delinquency, was no doubt one of the first to make this observation (1927). Delinquents have been found to occupy extreme positions on the intellectual distribution more frequently than chance expectancy (9). Also, clinical observation has shown the delinquent to be atypical in other ways. Not all, or even a significant number of delinquents deviate on the same dimensions or in the same direction. For example, some are extreme extraverts, while others are profoundly introverted. Some are wringing with anxiety while others are calm and "cool." Even on physical dimensions this phenomenon of extreme idiosyncrasy appears. Many delinquents are obese and many below weight. The physical development of many delinquents is premature while in others it is retarded. It is because of these great many ways in which delinquents are different, that a satisfactory typology of delinquency has not yet been developed. That multiple variation from the mean, without regard for dimension, is characteristic of delinquency was suggested by Pierson and Kelly who reported that 10 personality

* Received in the Editorial Office on June 17, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ Green Hill School and Fort Worden School are institutions of the Bureau of Juvenile Rehabilitation, Department of Institutions, Washington.

factors within a delinquent population were found to be very significantly different from control normals (7, 8). A single score was needed to express the gross deviation of an individual's personality from the mean. Such a measurement, called the Index of Idiosyncrasy, was then developed.

This study, therefore, investigated the relationship of anxiety, extraversion and general idiosyncrasy of personality to delinquent behavior, and this idiosyncrasy will be referred to as the Index of Idiosyncrasy.

B. METHOD

The HSPQ was administered routinely to all male juveniles, aged 14 years and over, received by the Washington Bureau of Juvenile Rehabilitation from the courts between April 1961 and September 1962. Procedures of administration are described elsewhere; also, it was shown that the state-wide population of institutionalized delinquents appeared representative of a large section of the United States, exclusive of the megapolitan areas (7, 8).

The data were prepared for processing by the Pacific Northwest Computer Laboratory's IBM 709 research computer. Programs were written to compute the second-order factors of anxiety and extraversion described by Cattell (3, 4, 5).

DELST and the Index of Idiosyncrasy were also calculated. This index is defined as the sum of the absolute differences between the subject's factor score and the standardization population mean factor score, divided by the number of factors of the test. The formula follows:

$$II = \frac{\sum |X_f - M_f|}{N_f}$$

where: X_f = the individual's score on a personality factor, and

M_f = the population mean for the factor, and

N_f = the number of factors in the test.

Pearson product moment coefficients of correlation were then computed between each of the measurements. Table 1 shows these intercorrelations.

C. RESULTS AND DISCUSSION

One centroid factor was extracted from the matrix of intercorrelations using estimated communalities by the method described by Fruchter (6). A comparison of estimated communalities and obtained communalities revealed differences sufficiently large to necessitate a second- and third-factor extraction. The factor loadings thus obtained are shown in Table 1, together with the obtained communalities. Variable A (see Table 1) is the institution

number used for administrative purposes which ranges from one to nine. It was included as a dummy variable to provide ground for the figure of this small matrix. It will be noted that correlations with Variable A, as well as the loading on centroid factor I, are not significantly different from zero.

TABLE 1
MATRIX OF INTERCORRELATION

| | N = 850* | | | | |
|--------|----------|------|-----|-----|-----|
| | A | B | C | D | E |
| A** | — | | | | |
| B | -.04 | — | | | |
| C | .05 | -.57 | — | | |
| D | -.02 | -.78 | .55 | — | |
| E | .04 | -.30 | .12 | .32 | — |
| I*** | .00 | -.47 | .67 | .94 | .32 |
| h2**** | .00 | .22 | .44 | .88 | .10 |

* Where $N = 850$ and $r = .09$, $\text{Sign.} = .01$.

** A = Dummy variable.

B = Second-order factor of anxiety.

C = Second-order factor of extraversion.

D = DELST delinquency potential (see text).

E = Index of Idiosyncrasy (see text).

*** Loadings on centroid factor I.

**** Communalities.

DELST is seen to be negatively correlated with anxiety at a high level of statistical significance. There has been little agreement among workers in the field of delinquency regarding the relationship of delinquency to anxiety. Those familiar with the problem are aware of the voluminous literature dealing with the subject. Consider, as an example, the opinion of Bloch who states that the behavior of delinquents is better understood when it is recognized that it is "anxiety driven" (1). According to Bloch, "Substantial data (sic!) in this and in other cases indicate that the smooth and unruffled façade of the most severely psychopathic delinquent masks considerable and potentially overwhelming anxiety" (1). The present findings should help to clear the air on this matter. With a negative correlation of .78 ($N = 850$) between the second-order factor of anxiety and DELST there can be little doubt that the anxiety level of this delinquent population is significantly below normal.

Table 1 further shows a high positive correlation of .55 ($N = 850$) between DELST and extraversion. This is consistent with earlier findings in which HSPQ factors A and H were found to be very significantly above average in teenage delinquents, describing a frivolous, carefree, devil-may-care extraversion (7, 8). "Exaggerated indifference" aptly describes this syndrome of low anxiety and high extraversion.

DELST and the Index of Idiosyncrasy were found to be positively and very significantly correlated. This supports the many clinical observations mentioned above in which delinquents were observed to deviate from the norms in a great variety of ways. The relationship of these measurements then shows the delinquent to exhibit a very significant overall personality idiosyncrasy and to be extraverted and nonanxious.

DELST is loaded .94 on centroid factor I. This correlation is sufficiently high to suggest that DELST and Factor I are essentially the same entity, and DELST is established as a dimension of personality within this population. Further, the factor loadings show the same low anxiety, high extraversion, high Index of Idiosyncrasy pattern to exist. A study of the relationship of this dimension to other life-record data, extraneous to the measuring instrument, is needed.

These findings seem to raise some interesting theoretical issues. Simple difference from the mean on personality factors has been demonstrated to be significantly related to delinquency. Clinical observation of delinquents strongly suggests that other human dimensions, if fed into the Index of Idiosyncrasy formula, would distinguish delinquents from nondelinquents. For example, as if to punctuate his abhorrence of the mean, the delinquent has invented his own idiosyncratic hair style which he even projects on to drawings of the human figure. It might be speculated that American culture, in particular its public education system, has devoted most of its attention to the mean and developed programs for that numerically superior average, so that the deviant is driven to delinquency in search of personal and group identity. Regardless of the cause, those responsible for the care and treatment of adjudged delinquents have this phenomenon of high Index of Idiosyncrasy to work with.

There is, of course, nothing in the present findings that would lead to the conclusion that a high Index of Idiosyncrasy is confined to the delinquent population. Undoubtedly highly talented and creative individuals would also score high on this index. No attempt was made in the present study to relate delinquency to other such groups. The immediate concern, rather, is with the implications for treatment of delinquency which are found in these data. First, the possibility of a more precise treatment theory presents itself. Quite obviously such is impossible as long as therapists have no clearly defined objectives on so fundamental a therapeutic issue as anxiety. The principal investigator recently heard a panel of five authorities on the treatment of juvenile delinquency discuss group psychotherapy with delinquents. Each panelist gave as a primary therapeutic objective the reduction of anxiety.

Today in the field of delinquency, entire institutional programs vary, depending on whether anxiety is considered to be high or low. Worse yet, this fundamental datum is often ignored in planning treatment for the delinquent.

Recidivism rates, although not entirely consistent across the nation, do seem to raise a question as to the adequacy of present treatment methods, which are based on extremely nebulous foundation theory. At best, treatment models developed for neurotics are used and such models are superimposed over the delinquent whether they fit or not. A treatment theory is needed which would provide therapeutic objectives and show treatment response through measurement. Such a theory might well indicate that an elevation of the anxiety level is essential to successful treatment of delinquents and that the methods used to reduce anxiety only serve to reinforce the pathology. Further study needs to be done to identify by personality measurement the types of delinquents that clinical observation has for so long recognized. Perhaps the most crucial advantage of a theory of delinquency treatment based on factorial measurement is that, while contributing a necessary portion of surplus meaning, it would also generate a wealth of testable hypotheses which are so vitally needed.

REFERENCES

1. BLOCH, D. D. Some concepts in the treatment of delinquency. *Casework Papers*, 1954.
2. BURT, C. L. *The Young Delinquent* (3rd rev. ed.) London: Univ. London Press, 1938.
3. CATTELL, R. B. *Personality and Motivation Structure and Measurement*. New York: World Book, 1957.
4. CATTELL, R. B., BELOFF, H., & COAN, R. W. *Handbook for the I.P.A.T. High School Personality Questionnaire (the HSPQ)* (2nd ed.) Champaign, Ill. (1602 Coronado Dr.): Inst. for Pers. & Abil. Testing, 1962.
5. CATTELL, R. B., & SCHIEFER, I. H. *The Meaning and Measurement of Neuroticism and Anxiety*. New York: Ronald Press, 1961.
6. FRUCHTER, B. *Introduction to Factor Analysis*. Princeton, N.J.: D. Van Nostrand, 1954.
7. PIERSON, G. R., & KELLY, R. F. HSPQ norms on a statewide population. *J. of Psychol.*, 1963, **56**, 185-192.
8. ———. B.J.R. Psychology Research Report No. 1, Green Hill School and Ft. Worden School, Washington, 1963.
9. TYLER, V., & KELLY, R. Cattell's HSPQ as a predictor of the behavior of institutionalized delinquents. Psychology Research Report No. 2, Fort Worden School, Port Townsend, Washington, 1963.

Green Hill School
Box 600
Chehalis, Washington

Sacramento County Probation
Department
Sacramento, California



THE EFFECT OF GROUP VERBAL SUGGESTION AND AGE ON THE PERCEPTION OF THE AMES TRAPEZOIDAL ILLUSION*¹

Department of Psychology, Marymount College, Tarrytown, New York

MOTHER JANE MARIE MCGEE, R.S.H.M.

A. INTRODUCTION

The essential feature of the Ames trapezoidal illusion is that, under specified conditions, a rotating trapezoid figure will be perceived by the viewer as oscillating back and forth rather than as rotating. Much of the recent research carried out in order to determine the factors influencing the perception of this illusion has focused upon elements of an objective nature. Investigations on physical properties of the stimulus itself as well as on parameters such as background intensity, exposure time, speed of rotation and viewing distance (5, 6, 8, 13) represent such research. These and other similar studies have defined the more or less significant role played by such objective variables in increasing or decreasing the perception of the illusion as well as pointing up the variability which exists within and between subjects from session to session in the perception of the illusion.

More recently, Cappone (1) studied the effect of verbal suggestion on the perception of the Ames illusion, a factor which, up to the present, has not received attention of a consequential nature. The experiment represented an attempt to bring into sharper focus elements other than those of a purely objective nature as influential in the perception of the trapezoidal illusion. Cappone based her assumption that suggestion would be effective in producing an increase in the perception of the illusion on the theory, presented by Sherif (11), that the introduction of suggestion into an ambiguous stimulus situation would re-define the ambiguity of the stimulus by furnishing a degree of external structure, and this in turn would provide a frame of reference for the subject, other than his own,

* Received in the Editorial Office on June 25, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ This study was submitted in partial fulfillment of the requirements for the degree of Bachelor of Arts in Psychology at Marymount College, Tarrytown, N.Y. The author wishes to express her sincere gratitude to Dr. M. LaVerne LaDriere for her careful guidance and continued inspiration as mentor of this dissertation and to Professor Richard T. Zegers, S.J. for his critical reading of the manuscript.

on which to base his judgments. The results of Cappone's study showed that eight out of eight of the experimental subjects whose data were subjected to statistical analysis perceived the illusion at a significantly greater frequency rate after positive verbal suggestion was administered.

Explorations such as this give rise to many questions whose implications must be thoroughly examined before an understanding of the factors involved in perception of illusory phenomena can be reached. One such question which has arisen as a result of Cappone's study is that concerned with whether the effect of positive verbal suggestion given on a group basis would operate in the same fashion as it did on an individual basis. This question arises because Cappone achieved her results by administering suggestion to each subject individually. A second question which presented itself was that of age differences in susceptibility to suggestion. In other words, would susceptibility to suggestion of the Ames trapezoidal illusion be more pronounced in older or in younger subjects? The general findings in the literature indicate that susceptibility to suggestion increases from birth to about age seven or nine after which time it decreases (2, 3, 4, 9, 10, 12).

The present research was designed as an attempt to seek answers to the specific questions presented above. Using Cappone's (1) results as a basis, it was hypothesized, therefore, that group verbal suggestion designed to increase perception of the Ames trapezoidal illusion would effect a significant increase in frequency rate of the reversal phenomenon for two groups of female subjects, consisting of high-school freshmen and college students, naive to the experimental situation. It was further hypothesized that the high-school group would be more susceptible to suggestion and perceive a greater number of illusions after introduction of the independent variable than would the college group.

B. APPARATUS AND PROCEDURE

This investigation utilized an adaptation of the Ames trapezoid which was mounted on a shaft geared to a constant speed motor so that it completed 35 revolutions per minute. The target contained six windows and measured 13-1/2 inches in length with long and short edges of 12-3/4 and seven inches respectively. Luminous paint was used to coat the figure and it was charged by rotation for two minutes under direct illumination of a 100-watt bulb immediately before each experimental session. This permitted viewing in a completely darkened room free from background cues. Chairs were arranged in a semicircle so that all subjects were at a viewing distance of 10 feet from the target.

Fifty-four female subjects were divided evenly into three groups. Group I consisted of high-school freshmen with an age range of 13 to 15 years. Groups II and III comprised college students ranging from 18 to 21 years. All subjects had either normal vision or wore corrective lenses throughout experimentation. Groups I and II were the experimental groups. For them the first three sessions comprised the nonsuggestion period and the fourth and final session the suggestion period. The nine subjects in each experimental group who perceived the fewest number of illusions during Session 3 were selected to be recipients of the suggestion. Low perceivers were chosen because there is a maximum number of reversals possible, based on the speed at which the apparatus rotates, and it was assumed that some of the high perceivers would already have perceived the maximum prior to suggestion. Moreover, it was felt that subjects who perceived fewer reversals would be more "reality oriented" and, thus, less susceptible to suggestion. Group III simply viewed the rotating trapezoid in all four sessions but did not receive suggestion. The latter group was established to provide a differential baseline against which the effects of suggestion in the two experimental groups could be measured.

The procedure was arranged so that each group was tested separately for the four sessions which were spaced consecutively at 24-hour intervals. The first two sessions were practice sessions in order to familiarize the subjects with the experimental situation. Sessions 3 and 4 were the sessions critical to this experiment since any significant difference between the results of the experimental subjects in these two sessions would indicate the operation of suggestion. An individual session consisted of a dark-adaptation period of seven minutes, one practice trial and 10 test trials each of which was three minutes in length with a one-minute rest period between each trial. Total session length was approximately an hour.

Prior to the first session, subjects were given a short instruction on the apparatus used and a description of the reversal phenomenon. Each subject was provided with a pencil and a small pad at the beginning of each session with which to record reversals perceived, and given opportunity to practice marking them without looking as they would be required to do this during actual sessions. Subjects were asked to watch the revolving apparatus during the three-minute periods and to mark their trial pages with a small "one" each time they perceived a reversal. The total number of reversals perceived by each subject in Groups I and II during Session 3 were ranked to separate the low perceivers from the high perceivers. Immediately prior to Session 4, the names of the high and low perceivers were called out and

the nine subjects who had perceived the highest number of reversals during Session 3 were commended and told to continue exactly as before. It was then made clear to the remaining nine subjects that their results showed that they had perceived too few reversals for their age group and that they would have to make greater efforts at cooperation and concentration if they were to achieve satisfactory results. Since the purpose of the study was to "group suggest" to low perceivers that they should be seeing a greater number of reversals, this step in the procedure was done publicly so that all subjects were aware of who had done "well" or "badly" up to that point. Session 4 was then carried out exactly as the previous three.

C. RESULTS AND DISCUSSION

In order to ascertain the overall effects of suggestion and nonsuggestion on the number of illusory reversals perceived, the data obtained from all subjects in the experimental and criterion groups during Sessions 2, 3 and 4 were submitted to analyses of variance. The data of Session 1 were not treated statistically since it was thought that considerable learning due to adjustment to the experimental situation was involved in this session.

The first analysis was performed on the data of all subjects in the experimental and criterion groups for nonsuggestion Sessions 2 and 3. The results are presented in Table 1 where it can be seen that the differences between groups and sessions are insignificant. This fact gives reasonable

TABLE 1
AN ANALYSIS OF VARIANCE COMPARING PRESUGGESTION SESSIONS 2 AND 3 FOR
GROUPS I, II AND III

| Source | SS | df | MS | F | Sig. |
|---------------------|-----------|----|---------|------|------|
| Subjects: Group I | 1,580,108 | 8 | 224,193 | — | — |
| Group II | 1,223,075 | 8 | | — | — |
| Group III | 2,577,454 | 8 | | — | — |
| Groups | 448,943 | 2 | 224,471 | 1.00 | — |
| Sessions | 771 | 1 | 771 | 0.02 | — |
| Subject vs. session | 1,027,423 | 24 | 42,809 | — | — |
| Session vs. age | 10,446 | 2 | 5,223 | 0.12 | — |
| Total | 6,868,220 | 53 | — | — | — |

evidence for two conclusions: First, justification of the criterion group used, since none of the groups differed significantly in the number of illusions perceived under conditions of nonsuggestion. Second, since the difference between Sessions 2 and 3 is insignificant, any further learning process after Session 1 had been completed would seem to be exerting an insignificant influence. The results seem to indicate, therefore, that continued exposure

to the stimulus situation did not, of itself, cause an increase in reversals perceived. This is in accord with the results of Cappone (1) who also found that a practice effect did not occur under conditions of nonsuggestion. The results also concur with those of Murray (7) who found that learning was completed during the first hour of viewing.

The data obtained from the criterion group were further studied by performing a second analysis of variance on Sessions 3 and 4. The purpose of the analysis was to ascertain whether the number of reversals perceived by this group during Session 4 was comparable to the number perceived during Session 3. That there was no significant increase in illusions perceived between the two sessions is evident from the *F* ratio presented in Table 2.

TABLE 2
AN ANALYSIS OF VARIANCE COMPARING NONSUGGESTION SESSIONS 3 AND 4 FOR GROUP III

| Source | <i>SS</i> | <i>df</i> | <i>MS</i> | <i>F</i> | Sig. |
|----------|-----------|-----------|-----------|----------|------|
| Sessions | 110,293 | 1 | 110,293 | .42 | — |
| Subjects | 4,203,641 | 16 | 262,727 | — | — |
| Total | 4,313,934 | 17 | — | — | — |

These results again point up the fact that continued exposure to viewing the apparatus was not, in the conditions set up for this investigation, the cause for an increase in perception of the reversal phenomenon.

The discussion so far has presented two conclusions. The first of these is that the number of illusions perceived during nonsuggestion Sessions 2 and 3 was approximately the same for all groups. The second is that the criterion group, which was not exposed to suggestion prior to Session 4, continued to perceive the illusion at the same frequency rate during this session as it had during the previous sessions. The implications of these facts become apparent when related directly to a discussion of the results of the final and critical analysis which was performed on the data of the two experimental groups for Sessions 3 and 4. As shown in Table 3, the analysis performed on these data revealed a difference between the pre- and postsuggestion conditions which was significant at the .01 level of confidence. These data point to the positive influence of verbal suggestion in effecting a significant increase in the number of illusory reversals perceived in Session 4 and are in direct agreement with Cappone's (1) results. It would seem reasonable to assume that had suggestion not been administered, the experimental groups would have continued perceiving the illusion at the same frequency rate as in Sessions 2 and 3. The basis for this assumption lies in the fact that there was no significant difference between any of

TABLE 3
AN ANALYSIS OF VARIANCE COMPARING PRE- AND POSTSUGGESTION SESSIONS 3 AND 4 FOR
GROUPS I AND II

| Source | SS | df | MS | F | Sig. |
|---------------------|-----------|----|-----------|-------|------|
| Subjects: Group I | 1,215,973 | 8 | 153,087 | — | — |
| Group II | 1,233,419 | 8 | | — | — |
| Groups | 979,770 | 1 | 979,770 | 6.40 | .05 |
| Sessions | 2,693,428 | 1 | 2,693,428 | 41.00 | .01 |
| Subject vs. session | 1,054,719 | 16 | 65,920 | — | — |
| Session vs. age | 181,618 | 1 | 181,618 | 2.76 | — |
| Total | 7,358,927 | 35 | — | — | — |

the sessions in the number of illusions perceived by the criterion group (Tables 1 and 2) and also in the fact that the two experimental groups were perceiving the illusion at the same frequency rate during Sessions 2 and 3. The results also substantiate the view taken by Sherif (11) that suggestion seems to operate as a modifying power in an ambiguous stimulus situation by creating an external frame of reference on which the subject can base his judgments. In conclusion, the data support the hypothesis that verbal suggestion given to a group operates in the same manner as does suggestion given individually.

As to the question of age differences in susceptibility to suggestion, the hypothesis had predicted a greater increase in the perception of the illusion for the younger group after administration of suggestion. Reference to Table 3 reveals that the difference between the two groups was significant at the .05 level of confidence with the college group scoring relatively higher than the high-school group in Session 4. These results disagree with the general findings in the literature that suggestion decreases with age (2, 3, 4, 9, 10, 12).

D. SUMMARY AND CONCLUSIONS

This investigation was concerned with the influence of group verbal suggestion in effecting a significant increase in perception of the Ames trapezoidal illusion and with the question of age difference in susceptibility to suggestion.

The results discussed in the report can be summarized as follows:

1. Group verbal suggestion effected a significant increase in perception of the Ames illusion at the .01 level of confidence for the two experimental groups. This finding is in direct agreement with the work of Cappone (1).
2. Disagreement with the general findings in the literature that susceptibility to suggestion decreases with age (2, 3, 4, 9, 10, 12) was shown

in the fact that the greater number of illusions perceived by the college group during the suggestion session was significantly different at the .05 level of confidence from the number perceived by the high-school group.

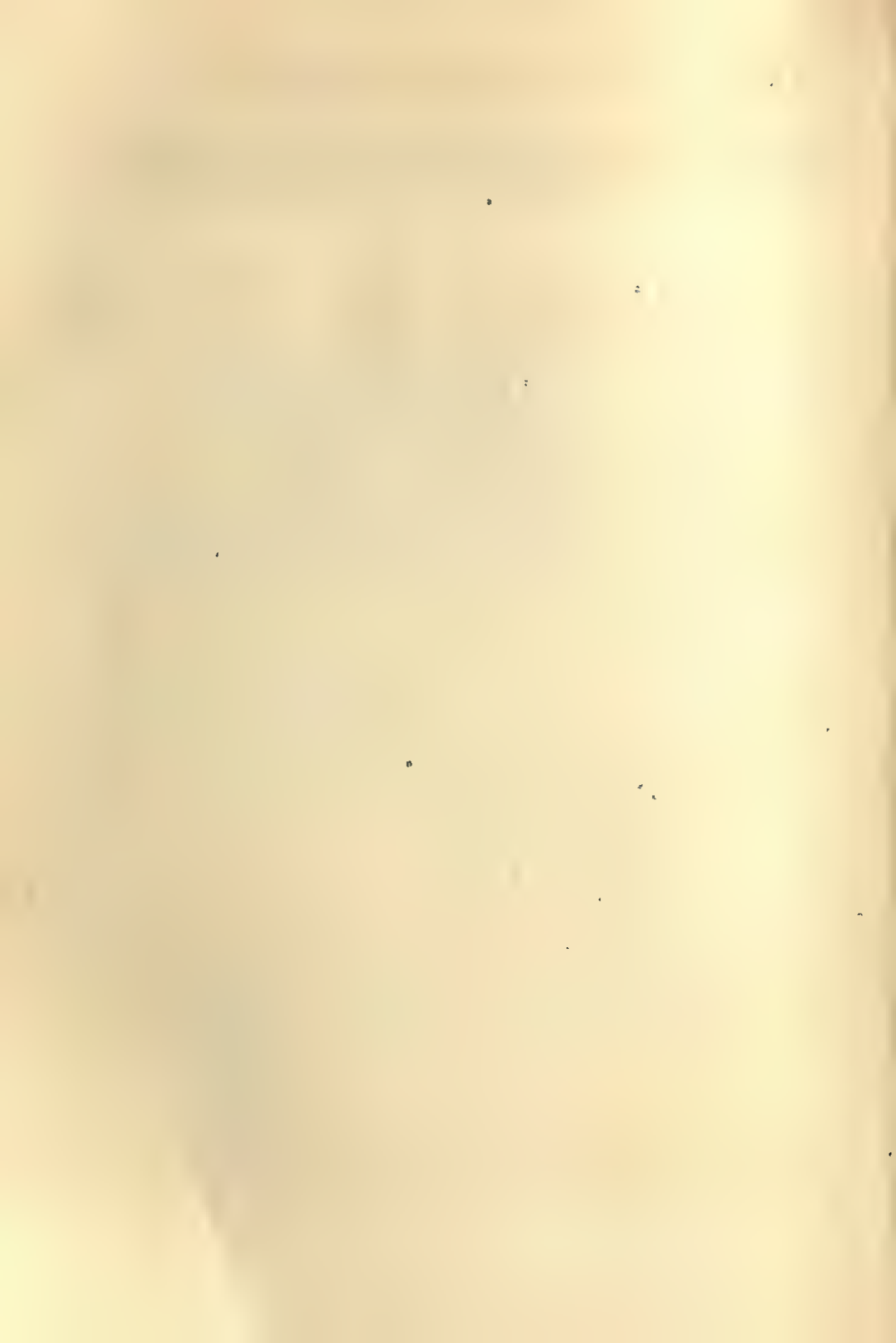
REFERENCES

1. CAPPONE, M. K. The effect of verbal suggestion on the reversal rate of the Ames trapezoidal illusion. Unpublished Master's dissertation, Fordham University, New York, 1963.
2. GILBERT, J. A. Researches on the mental and physical development of school children. *Stud. Yale Psychol. Lab.*, 1894, **2**, 40-100.
3. LODGE, J. R. The Illusion of Warmth Test for suggestibility. *Forum Educ.*, 1926, **40**, 180-186.
4. MESSERSCHMIDT, R. The suggestibility of boys and girls between the ages of six and sixteen years. *J. Genet. Psychol.*, 1933, **43**, 422-437.
5. MULHOLLAND, T. B. Motion perceived while viewing rotating stimulus objects. *Amer. J. Psychol.*, 1956, **69**, 96-99.
6. ———. The "swinging disk" illusion. *Amer. J. Psychol.*, 1958, **71**, 375-382.
7. MURRAY, P. A. A quantitative study of the Ames oscillatory illusion. Unpublished Doctoral dissertation, Fordham University, New York, 1961.
8. PASTORE, N. Some remarks on the Ames oscillatory effect. *Psychol. Rev.*, 1952, **59**, 319-323.
9. REYMERT, M. L., & KOHN, H. A. An objective investigation of suggestibility. *Psychol. Bull.*, 1938, **35**, 683-684.
10. ———. An objective investigation of suggestibility. *Charac. & Personal.*, 1940, **9**, 44-48.
11. SHERIF, M. A study of some social factors in perception. *Arch. Psychol.*, 1935, **27**, No. 187.
12. SMALL, M. H. Suggestibility of school children. *Ped. Sem.*, 1896, **13**, 176-220.
13. ZEGERS, R. T., & MURRAY, P. A. Perception of distortion: An experimental approach to illusion. New York: U.S. Naval Training Device Center, 1962 (Technical Report NAVTRADEVCEEN 506-1).

Department of Psychology

Marymount College

Tarrytown-on-the-Hudson, New York



STUDIES IN THE NEUROPHYSIOLOGY OF LEARNING: VII.
BAR-PRESSING RATE FOR CRANIAL SELF-STIMULATION
AS RELATED TO FREQUENCY OF RECTANGULAR
PULSES WHEN \dot{Q} IS HELD CONSTANT*¹

University of California, Los Angeles

J. A. GENDERELLI, RONALD D. PRIDDY,² AND JAMES R. AVERILL

A. INTRODUCTION

The discovery of the phenomenon of cerebral self-stimulation in the laboratory rat (20, 21, 23) was an event of capital scientific importance for learning theory. This finding shows that electrical stimulation of neurones in certain regions of the cerebrum has a positively reinforcing effect for the animal. The phenomenon frees the theorist from the necessity of involving himself in a number of ancillary considerations which serve to complicate and obscure the problems of learning and reinforcement. For in this case, the mere activation of an appropriately chosen set of neurones is sufficient to bring into being a state of affairs in the central nervous system which promotes the learning process; *viz.*, the consistent application of a prespecified stimulus, artificial as it is, has the effect of increasing the probability of occurrence of a behavior pattern. To invoke the stimulation of "pleasure centers" in the brain (22) to explain the phenomenon of cranial self-stimulation is a conceptual maneuver which is not acceptable in the context of a neurophysiological theory of learning.

Consider a pair of parallel (or concentric) electrodes imbedded at some point P_1 in the cerebral mass, and let a single rectangular (monophasic) pulse of current be applied through them of intensity i_0 and duration t_k . If these two parameters have suitable values, at least one neurone will be made to fire and if a microelectrode is situated at some point P_2 , a spike will be observed (ignoring inhibitory neurones). Keeping the intensity of the current fixed, if the duration of the stimulus pulse is increased to the value t_m , a spike will be observed at points in addition to P_2 . This is to be expected by virtue of the fact that a sufficient value of the product $i \cdot t$ will reach the

* Received in the Editorial Office on June 27, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ Financial support from the Faculty Research Committee of the University of California, Los Angeles, is gratefully acknowledged.

² Deceased.

threshold of a greater number of neurones than before. If, with i still fixed, the duration of the stimulus pulse is increased further, the number of points where spikes are recorded will also increase, although the number of such active loci will rapidly reach some limit (10, 14).

Because of the lattice-like structure of the brain in many places, we may expect, further, that in many instances, as the duration of the stimulus pulse is increased, more than one spike will appear at P_2 as a consequence of a single pulse. The number of multiple spikes resulting from increased stimulus duration will also rapidly reach a limit. Analogous results are to be expected if the duration of the stimulus pulse is kept at the value t_k and the intensity of current is increased (6, 19).

It is thus seen that the size of the product $i \cdot t$ characteristic of the stimulus pulse applied at a fixed point in the cerebrum influences the degree of transsynaptic activity distal to that point.

Now, let i and t be held constant at the levels i_0 and t_k , respectively, and let the monophasic pulse be applied twice with some interval Δt between. For a suitable value $(\Delta t)_j$, we can expect two spikes at the point P_2 and, if a series of stimulus pulses separated by $(\Delta t)_j$ is given, we will obtain a corresponding series of spikes. Of even greater interest is the fact that if the stimulus is applied at a repetition rate determined by $(\Delta t)_j$, spikes will appear at points not previously activated by a single pulse and, further, that some points previously activated by a single pulse will now be silent (2, 4, 15, 17).

If we consider an element of volume ΔV in the cerebral mass as constituted of the dendrites and soma of a single neurone inclusive of the initial segment of its axon, we may define the degree of transsynaptic activation for that region at the time t as the rate of firing detected there by a microelectrode. This rate is determined primarily by the number of presynaptic impulses entering the region in unit time. Thus if the number of presynaptic impulses impinging per unit time upon the soma or dendrites of a given nerve element is zero, transsynaptic activation is zero (ignoring "spontaneous" discharges); as this number increases, the rate of postsynaptic discharge will increase also, very probably as some negatively accelerated function of the presynaptic rate. The presynaptic rate, in its turn, would be determined by the parameters of the stimulus applied to the presynaptic fibres.

The experiments to be described in the present report are guided by the hypothesis that the reinforcing effects of intracranial stimulation are some mathematical function of the degree of transsynaptic activation. The *locus* of transsynaptic activation is, of course, basic to the phenomenon but only in the sense that Area 17 is indispensable for the acquisition of visual dis-

criminations. Our purpose here is to determine whether, if a pattern of behavior occurs at all in consequence of a defined situation, it is influenceable by varying the degree of transsynaptic activity which is involved.

The hypothesis may be investigated with a minimum of ambiguity, perhaps, by studying the effects of varying repetition rates of rectangular pulses upon some simple behavior variable such as rate of lever pressing. In so doing, we must keep the other parameters of the stimulus constant, *viz.*, intensity, duration of single pulse and durations of impulse train. As regards duration of the rectangular pulse, the objection has been made that in keeping it constant, any variation in the repetition rate results in a variable amount of electricity (Q) applied to the cerebral elements being stimulated in unit time, and that the effects of stimulus repetition rate will be confounded with Q . To meet this criticism it will be necessary to adjust the duration of the stimulus pulse to the repetition rate so that the Q delivered to the nerve tissue in unit of time is equal for all repetition rates. In doing this, however, if the repetition rate is sufficiently great, the duration becomes so short that the *constant quantity relation* of nerve stimulation is not fulfilled (11, 24). Thus a single pulse will not stimulate, but will require the temporal summation of several pulses. This fact, in its turn, will lead to confounding since two different repetition rates may yield identical degrees of transsynaptic activation for reasons which are not germane to the hypothesis under investigation. If, therefore, the constancy of Q is to be observed for various repetition rates it will be necessary that in the highest repetition rate used, the associated duration of a single pulse should be such (*i.e.*, be sufficiently long) as to be capable, at the intensity used, to stimulate some neurones.³

B. METHOD

1. *Experimental Design*

The factorial manipulation of three independent variables dictated the design of the experiment. The variables were:

a. *Repetition rate of stimulus.* The stimuli used* were rectangular pulses from Grass (SD-5A) stimulators. Eight different frequencies were employed: 20, 40, 60, 80, 100, 125, 150 and 200 pps.

b. *Stimulus phase.* For each repetition rate both a monophasic and a biphasic form of the stimulus was used. A clean biphasic signal was achieved

³ Since the chronaxie of the nerve elements at the point of electrode placement is unknown, we can only surmise what this time value should be. A possible rule of thumb is that no pulse duration used, of whatever phase, should be less than one millisecond.

by utilizing two stimulators "in series" such that the first unit triggered the second and in such wise that no delay intervened between the two aspects of the signal.

c. *Permutation of repetition rates.* For each *S* the experiment lasted 16 days. On each of these days he was subjected to each of the eight repetition rates either in monophasic or biphasic form. Thus eight days were devoted to biphasic stimulations and eight, to monophasic. A latin-square table determined the eight permutations to be used. The same table was applied to both the monophasic and biphasic forms of the stimulus.

Table 1 gives the regimen followed by the *Ss* in the experiment. From the table it will be noted that a confounding between permutation of repetition rates and days of practice occurs in this design.

2. Experimental Procedure

The *Ss* were given preliminary training in pressing a bar for intracranial stimulation (sine waves, 60 cycles per second, .5-second stimulus train for each separate press). This training continued (usually for a period of two weeks) until the *Ss* were judged to have reached a stable pressing rate. During this period a suitable stimulus intensity was determined for each *S*. The range of current intensities was from 60 to 100 μ A. A monitoring oscilloscope was placed across a .5-megohm resistor in series with the stimulus leads. By adjusting the stimulus voltage, it was possible to keep the amplitude of the oscilloscope trace at a fixed height during the course of the experiment for any given animal. The high resistance placed in series in the system served to minimize the effects of polarization.

In order to realize a constancy of *Q* for all repetition rates, whether for biphasic or monophasic stimuli, the duration of each pulse was variously adjusted for the different repetition rates used. This schedule was designed to give a 50 per cent duty cycle to the stimulator(s) and is shown in Table 2.

Preliminary training finished, the first experimental sessions began the next day. All 16 experimental sessions conformed to the following pattern:

The animal was placed in the Skinner box and permitted to press for two minutes. The stimulus used was 60-cycle, and at the intensity required during preliminary training. At the end of this time, the 60-cycle stimulus was discontinued and the rectangular pulses applied with a delay of a fraction of a second (accomplished by flipping a single switch). The intensity of the stimulus from the Grass stimulator(s) had been previously adjusted to equal the rms voltage in the 60-cycle source. The *S* was permitted to press for a period of two minutes. At the end of this interval,

TABLE 1
PATTERNS OF PERMUTATIONS AND PHASES FOR DURATION OF THE EXPERIMENT

| Days | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|----------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| Series 1 | $P_1^{\phi_B}$ | $P_1^{\phi_M}$ | $P_2^{\phi_M}$ | $P_2^{\phi_B}$ | $P_3^{\phi_B}$ | $P_3^{\phi_M}$ | $P_4^{\phi_M}$ | $P_4^{\phi_B}$ | $P_5^{\phi_B}$ | $P_5^{\phi_M}$ | $P_6^{\phi_M}$ | $P_6^{\phi_B}$ | $P_7^{\phi_B}$ | $P_7^{\phi_M}$ | $P_8^{\phi_M}$ | $P_8^{\phi_B}$ |
| Series 2 | $P_1^{\phi_M}$ | $P_1^{\phi_B}$ | $P_2^{\phi_B}$ | $P_2^{\phi_M}$ | $P_3^{\phi_M}$ | $P_3^{\phi_B}$ | $P_4^{\phi_B}$ | $P_4^{\phi_M}$ | $P_5^{\phi_M}$ | $P_5^{\phi_B}$ | $P_6^{\phi_B}$ | $P_6^{\phi_M}$ | $P_7^{\phi_M}$ | $P_7^{\phi_B}$ | $P_8^{\phi_B}$ | $P_8^{\phi_M}$ |

Note: P represents a particular permutation; ϕ_B = biphasic, ϕ_M = monophasic, form of the stimulus.

Two of the S's were submitted to Series 1, and three to Series 2.

all power was turned off for a period of four minutes. Within 5-10 seconds of the end of the four-minute interval, the animal was given one or two "free shots" of 60-cycle stimulation. This had the effect of terminating the rest period and the *S* resumed pressing. The cycle was repeated exactly as before, with the exception that a different repetition rate of rectangular pulses was utilized, in accordance with the schedule appropriate to the *S* (see Table 1).

The division of each daily session into cycles consisting of the sequence 60 cycle, 2'—rectangular pulses, 2'—rest, 4', was dictated by the desire to avoid to some degree interaction between successive repetition rates. By giving *S* a four-minute "rest period" after each repetition rate and by beginning each new cycle with the same stimulus frequency (60 cycles), it was hoped that each successive repetition rate would impinge on the same "initial conditions."

Each daily session for a given *S* lasted 64 minutes (of this, 32 minutes was rest time) and was carried through without any interference on the part of *E*, who was invisible to *S* during the entire session. The circuitry of the apparatus system was such that all maneuvers such as "free shots," etc., were executed by suitable switching devices.

By means of counters forming part of the circuitry, records were available for the number of presses executed by *S* for each period. Cumulative timers indicated the amount of time during which the level was kept pressed down by the animal. By dividing this quantity for a given period by the number of bar presses, it was possible to obtain the average length of bar press for that period. This quantity has a certain interest since a given press yielded a fixed .5-second duration of current regardless of the duration of level depression.

3. Subjects, Electrode Placement and Problem Box

Five male rats, age 180-250 days, were used in the experiment. These were of the F_3 generation resulting from the original cross between a Sprague-Dawley and Long-Evans strain and were black in body color and eye pigmentation. These animals have proved to be in our experience especially docile and sturdy.

The electrode unit, electrodes, wire leads, etc., have been described in detail elsewhere (7). Electrodes were placed stereotaxically by means of the Krieg-Johnson apparatus and the stimulation target was the posterior hypothalamus. The same coordinates were used for the five animals, *viz.*, 4 mm behind the bregma, 1 mm to the right of the midline, and $8\frac{1}{2}$ -9 mm vertically down from the upper surface of the calvarium. No histological verification was made on the *Ss*, but was carried out on 21 animals in another

investigation where identical coordinate values were used. In this latter group of animals the electrode tips were verified to lie in the region of the posterior hypothalamus in the right hemisphere.

The bar-pressing apparatus consisted of a wooden enclosure 6" \times 10" \times 20". The top of the box was supplied with a mirror tipped at a 45 degree angle to permit observation of *S*. Through a slot near the bottom of one wall a plastic bar 3-3/4" wide was inserted for a depth of one inch. The bar was one inch from the floor and required a pressure of 25 gm through a distance of 1/8" to trigger the attached microswitch.

The animals were housed singly in home cages where food and water were available *ad libitum* at all times.

In all animals it was necessary to raise the value of the required current somewhere during the course of the experiment. This was done at the beginning of a daily session.

C. RESULTS

1. General Observations

All the *Ss* were "good pressers." The amount of current was so adjusted for each that no apparent somatic reaction to the electric shock was elicited. None of the animals suffered a "seizure" during the course of the experiment.

One item of interest is the behavior of the *Ss* during the four-minute rest period. *S* quite typically continued to press the bar several times after the current was cut off. The duration of these presses was longer than the preceding ones. As the animals became more practiced, however, the number of presses after the cessation of the current diminished sharply; so that after the first few days an *S* would press only once or twice after the power was turned off. Quite typically the animal would then retreat to the far side of the box and groom himself or just "wait." After the passage of several minutes—somewhere between 2½ and 3—the rat would approach the still "dead" bar and press once or twice; then retire once more. Sometimes there would be three or four presses in quick succession. In other instances the *S* might execute this maneuver two or three times during the course of the rest period. When the four-minute rest period was over and the *S* was given a "free shot" of 60-cycle voltage, the animal would very promptly hurry over to the lever and begin pressing; in some instances the rat would wander hurriedly around the floor of the box as if disoriented and seeking the location of the bar.

The behavior of the *Ss* to changes in the stimulus from 60-cycle sine wave to certain repetition rates of the rectangular pulses was most impressive.

TABLE 2
DURATION OF A SINGLE PULSE IN MILLISECONDS
(For all frequencies a 50 per cent duty cycle was maintained)

| Frequencies | 20 | 40 | 60 | 80 | 100 | 125 | 150 | 200 |
|-------------|-------|-------|------|------|------|------|------|------|
| Monophasic | 25.00 | 12.50 | 8.32 | 6.24 | 5.00 | 4.00 | 3.32 | 2.50 |
| Biphasic | 12.50 | 6.25 | 4.16 | 3.12 | 2.50 | 2.00 | 1.66 | 1.25 |

Within a second after the change, there would be a sharp drop in the rate of pressing, followed very shortly by complete cessation. The *S* would "turn his back" to the bar and retreat toward a corner; sometimes, halfway toward the far wall, the animal would turn around and press a few more times. This pattern was quite consistent with all animals and occurred predictably when repetition rates—whether bipolar or monopolar—of 20 or 40 per second were used.

2. Quantitative Data

a. Bar presses. The average number of bar presses over a two-minute period for each frequency, monophasic and biphasic, is given in Table 3 for each of the five animals used. Each average is based on eight values.⁴ The averages based upon results given by the five *Ss* are given in Figure 1. Both curves show an inflection at a frequency of 100 per second. This frequency seems to have the greatest potentiating effect of the eight frequencies used. Biphasic rectangular pulses yield faster pressing rates for each frequency used than do monophasic pulses.

In Table 4 are given the average number of presses for a two-minute interval for each permutation of frequencies used. From Table 1 it will be seen that Permutation 1 was used for the first two days of the experiment (monophasic and biphasic pulses, but not necessarily in that order for a given *S*); Permutation 2, for the next two days, etc. Table 5 gives the bar-pressing performance for each of the five rats used, averaged over the eight frequencies.

Since each test frequency was preceded by two minutes of 60-cycle stimulation, it was considered worthwhile to record the average number of presses per two-minute period during the course of the experiment for the sinusoidal form of the stimulus. In the interest of compactness these are given as grand averages for each successive pair of days. Since the duration of the experiment was 16 days, there are eight values. Each of the values given in Table 6 is based on 16 two-minute periods (eight observations per day, two days).

⁴ Except in one instance in which gear failure resulted in the elimination of the use of one of the seven permutations.

TABLE 3
AVERAGE NUMBER OF BAR PRESSES FOR TWO-MINUTE INTERVAL

| Frequencies | 20 | 40 | 60 | 80 | 100 | 125 | 150 | 200 | |
|-----------------|------|-------|-------|-------|-------|-------|-------|-------|-------------|
| Monophasic | 22.7 | 43.9 | 41.5 | 59.3 | 95.9 | 62.1 | 61.4 | 48.9 | # 604 |
| | 26.4 | 29.6 | 44.6 | 76.9 | 55.0 | 59.6 | 51.9 | 62.4 | # 608 |
| | 20.3 | 44.0 | 70.1 | 84.7 | 69.9 | 79.5 | 78.0 | 55.6 | # 609 |
| | 7.9 | 11.7 | 26.1 | 64.9 | 71.3 | 67.7 | 71.7 | 54.9 | # 611 |
| | 12.5 | 13.1 | 36.4 | 63.1 | 102.9 | 49.5 | 67.4 | 64.0 | # 614 |
| Average | 17.9 | 28.4 | 43.7 | 69.8 | 79.0 | 63.7 | 66.1 | 57.1 | G.M. = 53.2 |
| Biphasic | 39.6 | 113.3 | 131.0 | 150.0 | 139.0 | 143.6 | 109.6 | 87.9 | # 604 |
| | 51.0 | 99.1 | 124.6 | 122.1 | 111.5 | 110.5 | 110.4 | 106.3 | # 608 |
| | 20.7 | 68.0 | 87.0 | 84.1 | 79.3 | 76.8 | 77.4 | 58.6 | # 609 |
| | 19.4 | 38.8 | 67.3 | 119.3 | 128.1 | 131.1 | 122.5 | 105.6 | # 611 |
| | 18.6 | 75.5 | 117.0 | 127.5 | 159.4 | 147.0 | 141.8 | 102.3 | # 614 |
| Average | 29.9 | 78.9 | 105.4 | 120.6 | 123.5 | 121.8 | 112.3 | 92.1 | G.M. = 98.1 |
| Overall Average | 23.9 | 53.7 | 74.6 | 95.2 | 101.2 | 92.8 | 89.2 | 74.6 | |

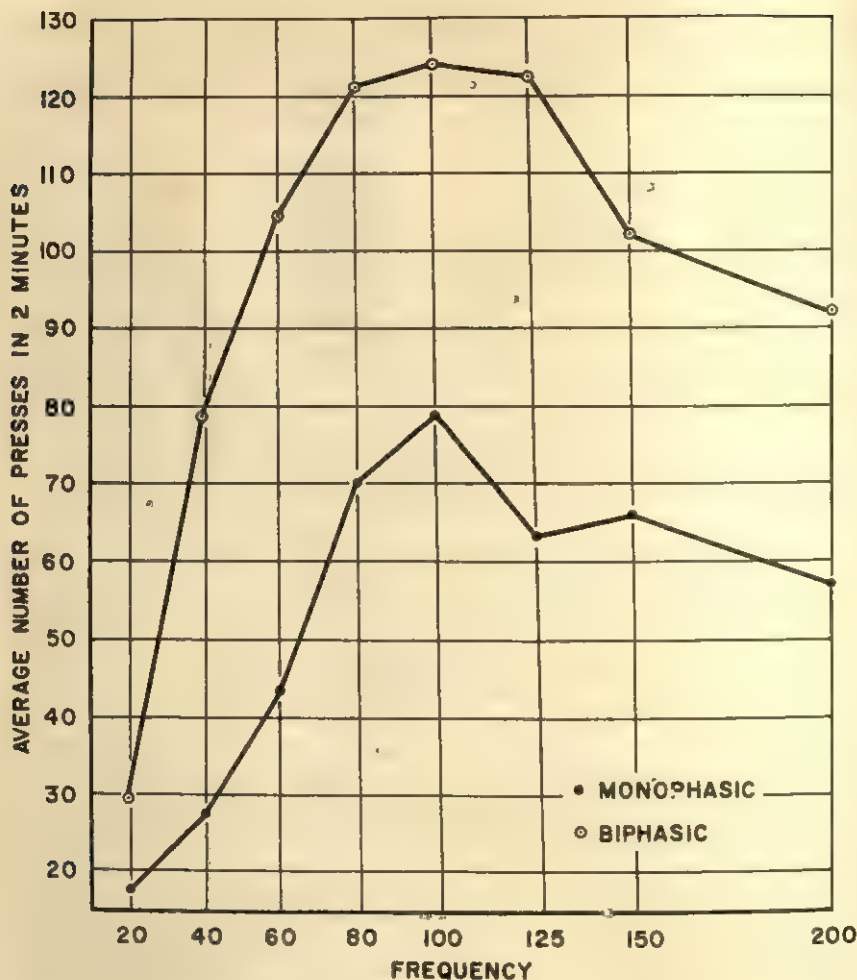


FIGURE 1

AVERAGE NUMBER OF BAR PASSES FOR A TWO-MINUTE PERIOD FOR VARIOUS REPETITION RATES

Clearly, the rate of pressing shows a downward trend during the course of the daily sessions.

The bar-pressing data generated by the eight frequencies of rectangular pulses was submitted to an analysis of variance. The results are shown in Table 7. The choice of suitable error terms for testing the main effects posed certain statistical difficulties because of the significance of several of

TABLE 4
AVERAGE NUMBER OF PRESSES PER TWO-MINUTE INTERVALS FOR EACH OF THE EIGHT PERMUTATIONS

| | P ₁ | P ₂ | P ₃ | P ₄ | P ₅ | P ₆ | P ₇ | P ₈ |
|----------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| ϕ_M | 79.0 | 61.0 | 42.6 | 48.2 | 46.3 | 49.7 | 45.8 | 53.2 |
| ϕ_B | 83.0 | 98.8 | 76.5 | 121.7 | 80.2 | 111.7 | 102.3 | 110.5 |
| Average | 82.0 | 79.9 | 59.5 | 84.9 | 63.3 | 80.7 | 74.0 | 81.9 |

TABLE 5
AVERAGE PERFORMANCES OF THE FIVE Ss FOR A TWO-MINUTE INTERVAL

| | #604 | #608 | #609 | #611 | #614 |
|----------|-------|-------|------|------|-------|
| ϕ_M | 54.4 | 50.8 | 62.8 | 47.0 | 51.1 |
| ϕ_B | 114.3 | 104.4 | 69.0 | 91.5 | 111.1 |
| Average | 84.3 | 77.6 | 65.9 | 69.3 | 81.1 |

TABLE 6
AVERAGE NUMBER OF PRESSES DURING A TWO-MINUTE PERIOD WITH 60-CYCLE STIMULATION
(Data given for each animal)

| Days | 1-2 | 3-4 | 5-6 | 7-8 | 9-10 | 11-12 | 13-14 | 15-16 |
|---------|-------|-------|-------|-------|-------|-------|-------|-------|
| #604 | 149.5 | 158.4 | 153.0 | 150.6 | 142.2 | 136.4 | 131.3 | 117.3 |
| #608 | 111.2 | 146.8 | 136.8 | 150.3 | 124.4 | 132.4 | 144.0 | 103.9 |
| #609 | 128.4 | 118.3 | 133.1 | 106.0 | 121.5 | 96.7 | 86.6 | 72.8 |
| #611 | 188.7 | 186.0 | 160.4 | 185.1 | 206.8 | 220.6 | 231.4 | — |
| #614 | 209.8 | 171.6 | 206.1 | 197.1 | 151.3 | 129.7 | 137.6 | 130.6 |
| Average | 157.5 | 156.2 | 157.9 | 157.8 | 149.2 | 143.2 | 146.2 | 106.2 |

Note: The average on the 15-16 days is probably reduced by the absence of #611.

the interactions. The treatment followed in these cases is described by Mood (16), and the *df* utilized for these composite error terms is calculated by means of the formula given by him (p. 348 *et seq.*).

The two dimensions of greatest experimental interest, namely, frequency and phase, are shown by the analysis to be significant factors: we may thus accept the trends shown in Table 3 as being real. The effect of permutation is negligible, as is individual differences among subjects (*cf.* Table 5). The interactions between frequency and phase and between frequency and subject are seen to be significant. The fact that the $f \times \phi$ interaction is significant poses difficulties of a statistical nature in carrying out some comparisons, since it is patently not a random variable. The triple interaction $P \times \phi \times S$ is also significant.

Since the frequency of electrical pulses is a significant factor in influencing the rate of bar pressing, it is of interest to determine whether there is a genuine inversion beyond 100 pps. This may be tested by computing a *t* based

TABLE 7
ANALYSIS OF VARIANCE OF AVERAGE NUMBER OF PASSES IN THE TWO-MINUTE INTERVAL
(f = Frequency; P = Permutation; ϕ = Phase; S = Subject)

| Source | df | SS | MS | F ratio | F |
|-----------------------------------|------------------|------------|------------|---|---------|
| f | 7 | 373,859.23 | 53,408.56 | $f/(f \times S) + (f \times \phi) - (f \times P \times S) \#1$ | 8.14*** |
| P | 7 | 48,988.48 | 6,998.35 | $P/P \times S$ | .51 |
| ϕ | 1 | 321,574.56 | 321,574.56 | $\phi/(f \times \phi) + (\phi \times S) - (f \times \phi \times S) \#2$ | 16.04** |
| S | 4 | 31,443.45 | 7,860.86 | $S/(P \times S) + (f \times S) - (f \times P \times S) \#3$ | .51 |
| $f \times P$ | 49 | 75,469.67 | 1,540.20 | $f \times P/f \times P \times S$ | 1.26 |
| $f \times \phi$ | 7 | 34,120.94 | 4,874.42 | $f \times \phi/f \times \phi \times S$ | 4.99** |
| $P \times \phi$ | 7 | 67,260.89 | 9,608.70 | $P \times \phi/P \times \phi \times S$ | .70 |
| $f \times S$ | 28 | 81,481.48 | 2,910.05 | $f \times S/f \times P \times \phi \times S$ | 1.91** |
| $P \times S$ | 28 | 384,213.98 | 13,721.93 | $P \times S/P \times \phi \times S$ | .99 |
| $\phi \times S$ | 4 | 64,757.27 | 16,189.32 | $\phi \times S/\phi \times P \times \phi \times S$ | 1.18 |
| $f \times P \times \phi$ | 49 | 72,082.21 | 1,471.07 | $f \times P \times \phi/f \times P \times \phi \times S$ | .97 |
| $f \times P \times S$ | 196 | 239,609.49 | 1,222.50 | $f \times P \times S/f \times P \times \phi \times S$ | .80 |
| $f \times \phi \times S$ | 28 | 26,964.60 | 963.02 | $f \times \phi \times S/f \times P \times \phi \times S$ | .63 |
| $P \times \phi \times S$ | 28 | 384,982.90 | 13,749.39 | $P \times \phi \times S/f \times P \times \phi \times S$ | 9.03*** |
| $f \times P \times \phi \times S$ | 188 ¹ | 286,387.63 | 1,523.34 | | |
| | 631 | | | | |

¹ Owing to apparatus failure, one animal missed one day of experimentation.

The formula used for calculating the df for the error terms is found in Mood (16, p. 384); #1, #2, and #3 = 12, 6, and 34 df , respectively.

** Significant at the .01 level.

*** Significant at the .001 level.

on the average presses with the 100 and 200 pps stimuli. As error term the $f \times S$ interaction was used. The means to be compared are shown in Table 3. For the biphasic stimulus each mean is based on 40 observations (five animals, eight observations for each frequency). For the monophasic stimuli, $N = 32$, since one session is missing. For the biphasic stimulation, $t = 2.6$, which is for 28 degrees of freedom significant at the .02 level. For the monophasic stimulation, $t = 1.8$ and is not significant. Thus for biphasic stimulation we obtain an easily identified inversion of effect: for monophasic stimulation the null hypothesis is not rejected in the form we have made it. However, the data suggest a downward trend between 100 pps and 200.

TABLE 8
MEAN TIME PER PRESS RECORDED IN SECONDS

| Frequency | 20 | 40 | 60 | 80 | 100 | 125 | 150 | 200 |
|------------|-----|-----|-----|-----|-----|-----|-----|-----|
| Monophasic | .80 | .71 | .64 | .53 | .51 | .60 | .58 | .58 |
| Biphasic | .57 | .55 | .50 | .52 | .51 | .51 | .57 | .62 |
| Mean | .66 | .59 | .54 | .52 | .51 | .54 | .58 | .60 |

b. Duration of lever press. Brain stimulation required depression of the lever, but in no case could a train longer than .5" be secured by the animal. Further stimulation required release of the lever and a new press. Cumulative time values were recorded and by dividing these quantities by the number of presses, an average duration of lever press was secured. In treating these data, the results of the eight permutations were lumped together, so that the average press duration for each frequency for a given phase is based on eight observations.

Results are shown in Table 8. It will be noted that for both monophasic and biphasic pulses the shortest press duration occurs at 100 pulses per second and increases on either side of this value, although the discrepancies are not great.

The average depression times were submitted to an analysis of variance, and the results shown in Table 9.

It will be seen that the impression given by Table 8 regarding the relationship between stimulus frequency and average duration of lever press is not borne out by the analysis. Nor is there a reliable difference due to phase or subjects. The interactions are all significant.

D. DISCUSSION

The present investigation was motivated by the hypothesis that the unit of function in the central nervous system, as in the peripheral, is the axon

TABLE 9

ANALYSIS OF VARIANCE OF THE AVERAGE DURATION OF A LEVER-PRESS DURING THE TWO-MINUTE TEST PERIOD
(Permutations have been pooled)

| Source | df | SS | MS | F ratio | F |
|--------------------------|----|----------|----------|--|---------|
| f | 7 | 3,613.80 | 516.26 | $f/(f \times \phi) + (f \times S) - (f \times \phi \times S)$ #1 | .99 |
| ϕ | 1 | 1,170.46 | 1,170.46 | $\phi/(f \times \phi) + (\phi \times S) - (f \times \phi \times S)$ #2 | 2.26 |
| S | 4 | 4,139.18 | 1,034.80 | $S/(f \times S) + (\phi \times S) - (f \times \phi \times S)$ #3 | 2.89 |
| $f \times \phi$ | 7 | 2,644.14 | 377.73 | $(f \times \phi)/(f \times \phi \times S)$ | 5.05*** |
| $f \times S$ | 28 | 6,117.82 | 218.49 | $(f \times S)/(f \times \phi \times S)$ | 2.92** |
| $\phi \times S$ | 4 | 857.41 | 214.35 | $(\phi \times S)/(f \times \phi \times S)$ | 2.86* |
| $f \times \phi \times S$ | 28 | 2,095.99 | 74.86 | | |

* Significant at the .05 level.

** Significant at the .01 level.

*** Significant at the .001 level.

The formula used for calculating the df for the error terms is found in Mood (16, p. 384); #1, #2, and #3 = 12, 8, and 10 df , respectively.

spike. For, while axon impulses are generated in the central nervous system by synaptic processes which are themselves continuous and nonquantized (1) and lead to discharges in the terminal arborizations of the axon, or telodendria (3), which, whether chemical or electrical in nature, do not bear a one-to-one relation to the axon spike, the spike itself is the basic communication "bit" in the nervous system. In attempting, therefore, to construct a rationale of nerve functioning as it relates to learning phenomena (or to any psychological phenomena, for that matter), the axon impulse is fundamental to our considerations. Since phenomena involving spikes are characterized by the parameter of frequency per unit time and duration of spike train (the amplitude of individual spikes may be neglected except for actuarial purposes), experiments dictated by the present hypothesis will be more preoccupied with the parameter of frequency in the physical stimulus than with the associated parameters of intensity and (in the context of rectangular pulses) with duration. Ward and Keesey, in their work with intracranial stimulation as it relates to stimulus parameters have concerned themselves with all three characteristics (12, 27), and have reported that in bar pressing for intercranial stimulation in the laboratory rat, the rate of bar pressing is influenced by altering, within limits, any one of the parameters. It has been pointed out by both investigators, however, that the influences of stimulus intensity and duration (when rectangular pulses are used) may be subsumed under a single factor, namely, the quantity of electricity, Q , contained in a single stimulus pulse. They were able to show that the rate of bar pressing was a positive—asymptotic (?)—function of the number of microcoulombs per pulse.

This certainly is an incomplete analysis of the situation insofar as the effective influence of stimulus frequency *per se* is concerned. Their findings are to be related to the well-known strength-duration function which applies to nerve stimulation. This function tells us to a first approximation that a fixed quantity of electricity Q is required to set up an impulse in a nerve fibre. However, in the context of the present problem we are concerned with the differential effects of the frequency of nerve impulses reaching a synaptic surface.

In the instance of electrodes implanted in the cerebral mass there will be electrical coupling with scores and probably hundreds of nerve elements with varying stimulation thresholds. If we confine ourselves to the stimulating potency of a single rectangular pulse, it is clear that as the quantity of electricity delivered by it is increased, an increasing number of nerve elements, up to some limit, is affected. Each of these elements will generate at least one axon spike: hence the extent of cerebral activation will vary as some positive asymptotic function of the quantity of electricity delivered by each stimulus pulse.

Since in the present experiment we are concerned with the frequency of axon spikes generated, we have attempted to eliminate quantity of electricity as a factor by making it constant. That is to say, the quantity of electricity delivered to the nerve elements per unit time is fixed for all frequencies, although, of course, the amount of electricity per stimulus pulse varies inversely with the repetition rate (*cf.* Table 2). In an earlier experiment (25), the amount of electricity was kept constant for each pulse, hence the total quantity per unit time varied linearly with the frequency used. By applying this restriction we inevitably courted the twin dangers of (a) polarization and nerve injury (13) for the slow repetition rates, and (b) of insufficient pulse duration at the high repetition rates. We attempted to defend against polarization by placing a very high resistance (500,000 ohms) in series with the stimulator output: this maneuver was quite successful, to judge from the tracings on the monitoring oscilloscope. As to (b), we set the duration of the pulse at 1.25 msec for the highest biphasic repetition rate (200). This duration, at 100 μ A, produces .125 microcoulombs per pulse. For the slowest frequency, 20 pps, a single biphasic pulse at the same intensity produces 1.25 microcoulombs.

In the absence of direct empirical evidence such as might be obtained from brain recordings at the time of stimulation, we cannot be certain that the average frequency of axon pulses associated with our stimulations at 200 pps is higher than for 20 pps. Indeed, it is not improbable that, although the

stimulus frequencies increased from 20 pps to 200 pps, at some point in the range the resulting spike frequencies at any arbitrary region of the cerebrum may have undergone a decrement. Since, for the biphasic stimulation at least, the bar-pressing rate at 200 pps is significantly lower than for 100 pps, such a state of affairs would be in accord with our hypothesis that degree of reinforcement, from a neurophysiological point of view, is a direct function of transsynaptic activation. A categorical verdict on the hypothesis, however, requires more conclusive evidence than that furnished by the present investigation.

The inversion shown in the curves of Figure 1, in any case, calls to mind the interesting findings of Coppée (5) and their mathematical treatment by Hill (8), to the effect that in stimulating a nerve fibre by a sinusoidal electric current, there is a specific frequency, characteristic of each nerve fibre, for which the required current is least.

Recent work (9, 26) has questioned the value of bar-pressing rate as a valid index of the "motivating force" created by intracranial stimulation. It is difficult to evaluate these findings, since the behavioral patterns in the experiments are so complex that they may involve a number of ancillary variables which may or may not be germane to the problem. Furthermore, it is unlikely that a stimulating factor will be related by the same mathematical function to different behavior variables.

E. SUMMARY

1. Bar pressing for cranial self-stimulation was studied in five rats which were subjected to rectangular electrical pulses at frequencies of 20, 40, 60, 80, 100, 125, 150 and 200 pps. Both monophasic and biphasic pulses were used. A factorial experimental design in which the dimensions were (*a*) frequency, (*b*) phase, (*c*) order of frequency presentations and (*d*) subjects, was utilized. The duration of rectangular pulses for each frequency was such as to make the amount of electricity delivered to S's brain per unit time the same for all frequencies and for monophasic and biphasic pulses.

2. It was found that the rate of bar pressing over a two-minute interval was a curvilinear function of the repetition rate of stimulation, with an inflection point at 100 pps. The general form of the function held for both monophasic and biphasic stimuli. There were no systematic effects due to the order in which the eight frequencies were presented in a given session, and there were no significant individual differences among the five animals. However, phase provided significant differential effects, biphasic impulses giving uniformly higher pressing rates than monophasic. There were several significant interactions.

REFERENCES

1. BODIAN, D. The generalized vertebrate neurone. *Science*, 1962, **137**, 323-326.
2. BROWN, M. L., & BRACKETT, C. E. Subcortical threshold voltages as a function of sine-wave frequencies. *J. Neurophysiol.*, 1958, **21**, 279-291.
3. BULLOCK, T. H. Neuron doctrine and electrophysiology. *Science*, 1959, **129**, 997-1002.
4. BURNS, K. D., & SMITH, G. K. Transmission of information in the unanesthetized cat's isolated forebrain. *J. Physiol.*, 1962, **164**, 238-251.
5. COPPÉE, G. La pararesonance dans l'excitation par les courants alternatifs sinusoïdaux. *Arch. Internat. Physiol.*, 1934, **40**, 1-58.
6. CURE, C., & RASMUSSEN, T. Effects of altering the parameters of electrical stimulating currents upon motor responses from the precentral gyrus of Macacca Mulatta. *Brain*, 1954, **77**, 18-33.
7. GENDERELLI, J. A., PROCTOR, D. M., & WOSKOW, M. H. Studies in the neurophysiology of learning: V. Differential effects of various rates of cerebral "driving" on behavioral modification. *Psychol. Rep.*, 1960, **7**, 123-141.
8. HILL, A. V. Excitation and accommodation in nerve. *Proc. Roy. Soc. [Biol.]*, 1936, **119**, 305-355.
9. HODOS, W., & VALENSTEIN, E. S. An evaluation of response rate as a measure of rewarding intracranial stimulation. *J. Comp. & Physiol. Psychol.*, 1962, **55**, 80-84.
10. KANDEL, E. R., SPENCER, W. A., & BRINLEY, F. J. Electrophysiology of hippocampal neurones: I. Sequential invasion and synaptic organization. *J. Neurophysiol.*, 1961, **24**, 225-242.
11. KATZ, B. *Electric Excitation of Nerve*. London: Oxford Univ. Press, 1939.
12. KEESEY, R. E. The relation between pulse frequency, intensity and duration, and the rate of responding for intracranial stimulation. *J. Comp. & Physiol. Psychol.*, 1962, **55**, 671-679.
13. LILLY, J. C. Injury and excitation by electric currents. In *Electrical Stimulation of the Brain*, D. E. Sheer, Ed. Austin: Univ. Texas Press, 1961.
14. MCCULLOCK, W. S. The functional organization of the cerebral cortex. *Physiol. Rev.*, 1944, **24**, 390-407.
15. MILHAIOVIC, L., & DELGADO, J. M. R. Electrical stimulation of monkey brain with various frequencies and pulse durations. *J. Neurophysiol.*, 1956, **19**, 21-36.
16. MOOD, A. M. *Introduction to the Theory of Statistics*. New York: McGraw-Hill, 1950.
17. MORUZZI, G. *Problems in Cerebellar Physiology*. Springfield: Charles C. Thomas, 1950.
18. NEWMAN, P. P. Single unit activity in the viscerosensory area of the cerebral cortex. *J. Physiol.*, 1962, **160**, 284-297.
19. NIELSON, H. C., DOTY, R. W., & RUTLEDGE, L. T. Motivational and perceptual aspects of subcortical stimulation in cats. *Amer. J. Physiol.*, 1958, **194**, 427-432.
20. OLDS, J. Runway and maze behavior controlled by basomedial forebrain stimulation in the rat. *J. Comp. & Physiol. Psychol.*, 1956, **49**, 507-512.
21. ———. Pleasure centers in the brain. *Sci. Amer.*, 1956, October.
22. ———. Differentiation of reward systems in the brain by self-stimulation techniques. In Ramey, E. R., & O'Doherty, D. S. (Eds.), *Electrical Studies on the Unanesthetized Brain*. New York: Hoeber, 1960.
23. OLDS, J., & MILNER, P. Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *J. Comp. & Physiol. Psychol.*, 1954, **47**, 419-427.

24. RUSHTON, W. A. H. Lapicque's canonical strength-duration curve. *J. Physiol.*, 1932, **74**, 424-440.
25. UYEDA, A. H., & GENCERELLI, J. A. Influences of rectangular pulses and sine waves of varying frequencies on brain self-stimulation in the laboratory rat. *Psychol. Rep.*, 1959, **5**, 641-647.
26. VALENSTEIN, E. S., & BEER, B. Reinforcing brain stimulation in competition with water reward and shock avoidance. *Science*, 1962, **137**, 1052-1054.
27. WARD, H. P. Stimulus factors in self-stimulation. *Amer. J. Physiol.*, 1959, **196**, 779-782.

Department of Psychology
University of California
Los Angeles 24, California

SEX DIFFERENCES AMONG SCHIZOPHRENICS IN THE INTERPRETATION OF THE HUMAN DIAD*¹

Psychological Laboratory, Wesleyan University

ROBERT H. KNAPP AND LINDA A. EHLINGER

A. INTRODUCTION

In a recent issue of the *Journal of Abnormal and Social Psychology* note was taken of the fact that it is an uncommon thing to report the sex of subjects used in psychological experiments. Beyond this, it was observed that studies of differences in the response of the two sexes to psychological tasks have been unduly neglected in the literature (1). The present article is devoted to a report of differences between male and female schizophrenics in response to a newly devised projective technique. It represents a variant and continuation of certain researches undertaken by Dr. Jules Holzberg and recently reported by him (2).

The test employed in this study has been previously reported by the author in two articles (3, 4). It is basically a device whereby the subject views two diadic silhouettes and is invited to indicate the nature of the social transaction occurring between them. It is postulated that the judgments required in this test reveal, among other things, the degree to which social relations are conceived of as fraught with friction or as characterized by amity, the degree to which the male or female figure is construed to be the initiator or the responder, and the degree to which different types of exchange are rarely or commonly perceived. Thus, it is hoped that this device will permit us to identify the style characterizing the social perceptions and interpretations of the subject.

The employment of a test of this type obviously offers two modes of interpretation. The first of these may be called psychophysical or *nomothetic*. Here our concern is not with individual differences but with the degree of consensus characterizing a population of individuals in the interpretation and evaluation of a stimulus. This particular approach was pursued in the two earlier articles. In the present study, on the other hand, we are concerned with the measurement of group differences in ascribing social

* Received in the Editorial Office on June 28, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ Support for this research was given by the National Institutes of Mental Health under Contract M-2178 to Wesleyan University.

transactions to the diadic figures described above. It may be said that this study is, therefore, concerned with ideographic rather than nomothetic objectives.

B. METHODS AND PROCEDURE

The subjects used in this experiment were 60 in all, comprising 30 male and 30 female chronic schizophrenics. They were selected from a schizophrenic population at a state hospital from which those too ill-oriented to comprehend the task were excluded.

1. *The Diadic Silhouette*

The Diadic Silhouette Test employed in this study has been reported elsewhere (3, 4). In the present instance certain simplifications were introduced in order to meet the limited capacities of our subjects and to simplify administration.

As employed in this study, the diadic silhouettes totalled 27. Each silhouette, showing the head and upper torso, presented a young man and young woman in profile confrontation. The 27 silhouettes were divided equally among three elevation patterns, namely (*a*) where the young woman occupied a higher vertical plane, (*b*) where they were equal, and (*c*) where the young man was higher. Each of the two figures was rotated in three degrees, namely backward, upright and forward. All combinations of these rotations were employed for each elevation pattern. Thus the total number of silhouettes involved three elevation patterns, three rotation degrees for the male figure, and three for the female, yielding 27 variants in all.

2. *The Subjects*

The subjects were shown the 27 projected silhouettes for approximately 45 seconds each with instructions to record the nature of the social exchange ensuing between the two figures. For this purpose a "menu" of possible transactions or exchanges which might provide an interpretation of each silhouette was supplied. This menu of transactions permitted the subject to designate which of the two persons represented was the initiator, which of four possible acts (telling, asking, offering, commanding) was involved, and whether or not the respondent reacted affirmatively or negatively in the exchange. It will be seen, therefore, that each subject was provided with 16 possible transactions from which he might select the one he considered most appropriate to each silhouette. It should be noted that each subject was required to respond only once to each of the 27 silhouettes but that he retained freedom of choice in his employment of the 16 items of the menu.

C. RESULTS

Our first results, presented in Table 1, show separately the incidence with which the two possible initiators, four possible acts and two types of responses are employed by male and female subjects respectively. It will be observed that there is a tendency, significant at about the six per cent level, for the male subjects to report female initiators and *vice versa*. Again, it will be observed that there is a significant difference between the employment of different acts by male and female subjects. More particularly, the female subjects prefer "telling" ($t = 2.56$; $p > .02$), the male subjects "commanding" ($t = 2.68$; $p > .01$). Finally, it will be observed that the male subjects are notably more inclined to employ propositions involving negative responses than are the females.

TABLE 1
THE INCIDENCE OF TRANSACTIONS INVOLVING POSSIBLE INITIATORS, ACTS AND RESPONSES
EMPLOYED BY MALE AND FEMALE SUBJECTS

| | Male subjects | Female subjects | χ^2 | p |
|-------------|---------------|-----------------|----------|-------|
| Initiator | | | | |
| He | 371 | 408 | 3.40 | > .06 |
| She | 437 | 400 | | |
| Act | | | | |
| Tells | 280 | 330 | 10.78 | > .02 |
| Asks | 172 | 184 | | |
| Offers | 126 | 111 | | |
| Commands | 230 | 183 | | |
| Response | | | | |
| Affirmative | 478 | 557 | 16.78 | > .01 |
| Negative | 330 | 251 | | |

Table 2 shows comparisons between our male and female subjects when the sex of the initiator and the four acts are considered together. For the acts "telling," "asking" and "offering" no significant relations may be adduced, but there is clearly a massive difference involving "command" transactions. Each group of subjects ascribes such exchanges to the figure of the opposite sex. This is notably true in the case of the female subjects who see commands issued by the male figure over twice as often as by the female figure.

Table 3 shows the incidence of responses classified according to the sex of the initiator and the type of response. It will be observed that male subjects show a significantly higher tendency than female subjects to ascribe negative responses regardless of the sex of the initiator.

Table 4 shows the frequency with which male and female subjects employ

TABLE 2
THE INCIDENCE OF RESPONSES BY MALE AND FEMALE SUBJECTS TO TRANSACTIONS
CLASSIFIED BY SEX OF INITIATOR AND CHARACTER OF ACT

| Initiator | Act | Male subjects | Female subjects | χ^2 | p |
|-----------|----------|---------------|-----------------|----------|-------|
| He | Tells | 133 | 154 | 0.04 | n.s. |
| She | Tells | 147 | 176 | | |
| He | Asks | 71 | 69 | 0.53 | n.s. |
| She | Asks | 101 | 115 | | |
| He | Offers | 59 | 56 | 0.30 | n.s. |
| She | Offers | 67 | 55 | | |
| He | Commands | 108 | 129 | 23.12 | > .01 |
| She | Commands | 122 | 54 | | |

TABLE 3
THE INCIDENCE OF RESPONSES BY MALE AND FEMALE SUBJECTS TO TRANSACTIONS
CLASSIFIED BY SEX OF INITIATOR AND CHARACTER OF RESPONSE

| Initiator | Response | Male subjects | Female subjects | χ^2 | p |
|-----------|-------------|---------------|-----------------|----------|-------|
| He | Affirmative | 219 | 281 | 8.16 | > .01 |
| He | Negative | 152 | 127 | | |
| She | Affirmative | 259 | 276 | 8.56 | > .01 |
| She | Negative | 178 | 124 | | |

transactions involving the four acts with positive and negative responses. Here we may observe that no significant differences resulted from comparing our two subject groups on "telling" and "offering" transactions. Transactions involving "asking," on the other hand, show that men subjects are more disposed than women to see this act followed by a negative response, a result standing at approximately the 4 per cent level of confidence. With respect

TABLE 4
THE INCIDENCE OF RESPONSES BY MALE AND FEMALE SUBJECTS TO TRANSACTIONS
CLASSIFIED BY NATURE OF ACT AND CHARACTER OF RESPONSE

| Nature of act | Response | Male subjects | Female subjects | χ^2 | p |
|---------------|-------------|---------------|-----------------|----------|-------|
| Tells | Affirmative | 220 | 269 | 0.83 | n.s. |
| Tells | Negative | 60 | 61 | | |
| Asks | Affirmative | 80 | 107 | 4.78 | > .04 |
| Asks | Negative | 92 | 77 | | |
| Offers | Affirmative | 72 | 68 | 0.40 | n.s. |
| Offers | Negative | 54 | 43 | | |
| Commands | Affirmative | 106 | 113 | 10.09 | > .01 |
| Commands | Negative | 124 | 70 | | |

to transactions involving "commanding," women subjects far more rarely than men see commands met with refusal.

The two remaining comparisons are perhaps worth reporting here. The first of these involved propositions in which the female figure is said to be "asking." Here we find that women significantly more often than men see requests initiated by the female figure met with compliance, as shown in Table 5. Finally, we may observe that when the male figure is seen as "commanding," women are more inclined, men less inclined, to see a compliant response.

TABLE 5
THE INCIDENCE OF RESPONSES BY MALE AND FEMALE SUBJECTS TO TRANSACTIONS INVOLVING "SHE ASKS" AND "HE COMMANDS"

| Initiator | Nature of act | Character of response | Male subjects | Female subjects | χ^2 | <i>p</i> |
|-----------|---------------|-----------------------|---------------|-----------------|----------|----------|
| She | Asks | Affirmative | 48 | 71 | 4.34 | > .05 |
| She | Asks | Negative | 53 | 44 | | |
| He | Commands | Affirmative | 50 | 83 | 7.76 | > .01 |
| He | Commands | Negative | 58 | 46 | | |

In summary, then, it would appear that male and female subjects differ in that men are (a) more inclined to see the female figure as the initiator, notably and significantly when "commanding" is the act involved; (b) more inclined to employ propositions involving "commanding," less likely to employ propositions involving "telling"; (c) more inclined to ascribe transactions with a negative response, especially for propositions involving "asking" and "commanding," notably when the female is the initiator.

D. DISCUSSION

This study represents the first attempt to report group or individual differences in performance on the Diadic Silhouette Test. We have dealt here only with the question of how often the various transactions were applied by our male and female schizophrenics to the 27 slides used as stimuli. Unexamined here is the question of possible differences in the interpretation of particular silhouettes by the two groups.

It is a clear limitation of this study that the subjects were confined to schizophrenics and the question naturally arises concerning a possible difference between their performance and that of normal males and females. Unfortunately, data bearing directly upon such differences are not available since the subjects in this experiment were given an abbreviated version of the original test with simplified instructions. Still, taking this into account,

some general observations are possible. Normal subjects appear generally to show a higher incidence of usage for the acts of asking and offering than do schizophrenics. The differences between male and female schizophrenics, further, do not seem generally sustained in the examination of normal subjects of the two sexes, especially in their attribution of the initiation of the exchange to the figure of the opposite sex and in the predilection of male schizophrenics for command and negative transactions. These observations, however, must stand merely as suggestions whose confirmation awaits the administration of a single common version of the test to comparable groups from normal and schizophrenic populations.

Some final comments should be made concerning the differences which we have obtained between male and female schizophrenics and their interpretation. In contrast to female schizophrenics, the males appear to view the female figure as the principal initiator, "telling" more rarely and "commanding" more commonly as the act involved, and most often a negative response, especially by the male respondent. There is contained in this pattern a virtual paradigm of the theoretic psychodynamics of male schizophrenia itself, in which the core of the male psychotic conflict is seen as revolving around resistance to the powerful demands of a female figure, presumably the mother. We see this with special clarity in the propensity of the male subjects to employ those particular transactions in which the female initiates requests or commands which are met with refusal.

The pattern of responses for the female schizophrenic is less clear of interpretation, but the strong predilection for the act of "telling" with its minimal affective commitment and the general disposition to see the male figure as initiator and the response as affirmative yields a picture quite different from that obtained among the males. It is congruent with the proposition that the female schizophrenic sees the world as male-dominated and one in which remote compliance is the preferred mode of adjustment.

REFERENCES

1. CARLSON, E. R., & CARLSON, R. A. E. Male and female subjects in personality research. *J. Abn. & Soc. Psychol.*, 1960, **61**, 482-83.
2. HOLZBERG, J. H. Sex differences in schizophrenia. Chapter in H. G. Beigel, *Recent Advances in Sex Research*. New York: Harper, in press.
3. KNAPP, R. H. Perceptual interpretation of the social diad: I. Judgments of acceptance. *J. Soc. Psychol.*, in press.
4. ———. Perceptual interpretation of the social diad: II. Character of relationship. *J. Soc. Psychol.*, in press.

Psychological Laboratory
Wesleyan University
Middletown, Connecticut

TRANSACTIONAL ASPECTS OF EGO THERAPY*

Western Psychological Center, Sherman Oaks, California

STEWART B. SHAPIRO

A. INTRODUCTION

In two previous papers (2, 3) the writer discussed a theory of ego pathology and introduced a therapeutic method designed to reduce this pathology. The purpose of this paper is to illustrate certain transactional aspects of what the writer has called "ego therapy." This paper will extend the ego-therapy approach to the kind of interpersonal transactions found in sensitivity training and various group therapies, including marital group therapy and family group therapy.

B. THEORETICAL PREMISES

It is a major theoretical premise of the writer that there is no such thing as a single "psychological self." There is, instead, a system of "subselves," located in the ego and more or less unified in a personal identity, depending on the strength of the synthesizing function of that ego. The structure of these subselves and their interrelations are of the utmost importance in determining behavior. The primary pathology in an adult personality "in trouble" is located within the ego itself, and the model for therapy of this ego pathology is a kind of family group therapy within the total ego organization. The emphasis is on training the patient to observe his family of subselves in action and to develop a conscious choice of which subselves he may wish to change and which he may want to "live with."

Transactional ego therapy involves the relationship between the "inner families" of several people. Analysis of these "interfamily" relationships often throws more light on each of the "internal families," and their internal games as well as the characteristic "interfamily (interpersonal)" relationships and games. Like Berne's (1) sequence of training in his transactional analysis, transactional ego therapy depends on first identifying and observing the structure of the personality and then repeatedly observing how one person's cast of subself characters interacts with other casts or how one cast interacts with the total group as a single generalized "other." It is as though the cast of

* Received in the Editorial Office on July 3, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

characters on one stage, absorbed in playing the roles assigned for a given play, were interrupted by the cast on a different but adjacent stage. Both plays stop, and a new play begins. The new play involves two or more people, two or more scripts and two or more inner casts. However, it is often remarkable how repetitive and similar the new play may be to each of the old "single" plays. The audience (other group members) may or may not approve of the "production."

There are a few theoretical premises in the ego-therapy approach which have special relevance to and underlie all the different technique models in transactional ego therapy. These premises are as follows:

1. Intrapersonal (subself) communication can be viewed as transactional. Not only are there identifiable subelves within the ego, but these subelves have "human relations" among themselves.

2. There is a marked similarity between intrapersonal (subself) transaction and interpersonal transaction. Both are transactions in the true sense of the term. Although not ordinarily conceptualized this way, subself communication has all the characteristics of transaction-sending and receiving of messages, meaningful motivated interchange, etc.

3. Nearly all social transactions and all patterned transactions have some intrapsychic meaning. One of the major goals of transactional ego therapy is to help the individual discover how he uses social transaction to play *internal* games. Another way of stating this principle is that external social games are designed and played for subself game strategy.

4. Transactional ego therapy helps to determine at what points a subself system can be reached and changed. Often social transaction is the process by which change is effected. Even the most structurally oriented individual ego therapy, after all, depends almost entirely on the social transaction between therapist and patient.

C. TECHNIQUE MODELS

The balance of this paper will be devoted to the listing of five of the many different technique models which can be derived from this theoretical approach. One of these will be illustrated in detail.

1. *The Mirror Model*

This model is taken from the mirror technique of psychodrama. Here, as in most of the other models, the patient's individual subself structure is established first. The main characters and their interrelationships are developed, and then gradually, other group members are sent in as doubles, so to speak,

ultimately replacing the original patient. Finally, as an outside observer, the patient sees his own characters playing out one of his typical subself conflicts. The doubles are chosen for the similarity to their own problems.

2. *The Empathy Model*

This was designed in a sensitivity training program in response to the question: How can one person, *A*, learn empathy for another person, *B*, whom *A* doesn't understand and feels is weak? In many cases a simple psychodramatic role reversal cannot accomplish this. ("I *can't* play the other guy—I just don't dig him!") The empathy model involves *A* in a series of role playing steps in *B*'s presence which bring out the identification of the rejected, weak child in *A*'s subself structure with *B*'s child, whom *A* perceives as weak. If successful, *A*'s child becomes an ally of *B*'s child. The presence of *B* is very important to test the validity of this alliance.

3. *The Eric Berne Model*

In this model, transactions from the subselves of one person to the subselves of another are analyzed as to whether they are complementary, crossed, etc. (e.g., Is the message which was sent to the adult answered by the adult or by the child?) In the ego-therapy approach the internal games are also stressed.

4. *The Group as a Subself*

In some subself systems the "generalized other" is incorporated as a well-defined active agent, usually representing society, a peer group or an audience, and often is an ally of the critical, evaluative, superego parental subselves. A transactional ego-therapy model of considerable help in social anxiety problems could be called the "group-as-a-subself" technique. In this model the subself structure of the patient is exposed and the whole group plays the audience subself. Interactions between the group as an audience subself with the other subselves of the patient is most instructive, especially when contrasted with the expectations and reactions of the patient's inner-audience subself. The mechanism of projection in a group situation can be very clearly demonstrated to the patient and the group by this procedure.

5. *The Marriage Transference Model*

The one model which will be presented in detail and illustrated by a protocol could be called the "Marriage Transference" model. Of the various transference reactions which occur in a marriage the game of "How dare you strike my child!" has been chosen to illustrate this model. The "marriage trans-

ference" model lends itself to division into the four successive steps listed in Table 1 and illustrated in Figure 1.

TABLE 1
THE "MARRIAGE TRANSFERENCE" MODEL

| | Training steps | Types of transaction |
|--------|---|--|
| Step 1 | Presenting the problem in interpersonal terms | A sample of everyday marital conflict |
| Step 2 | Training setup | Conversion of internal structure to intrapsychic transaction |
| Step 3 | The problem in subself terms | Subself interpersonal transaction |
| Step 4 | Observer training | Observation and analysis of model interpersonal transaction |

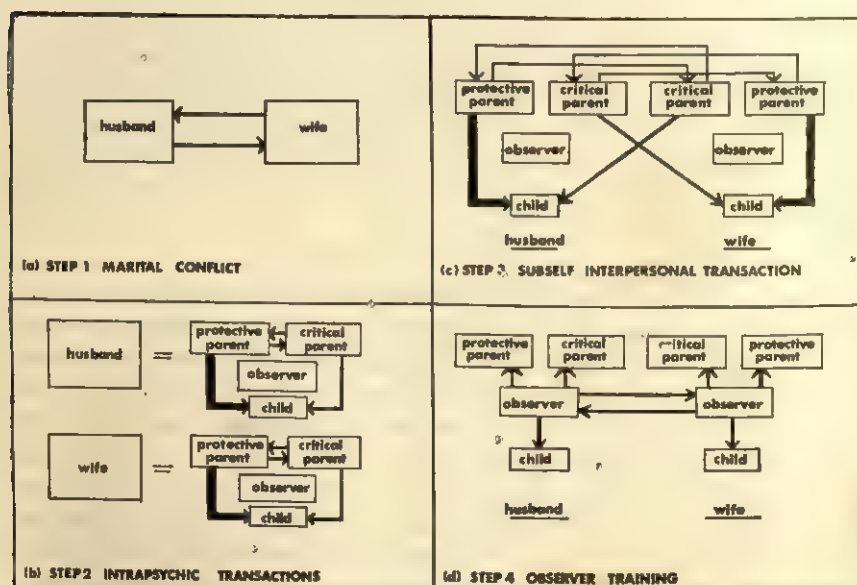


FIGURE 1
ILLUSTRATION OF THE MARRIAGE TRANSFERENCE MODEL

D. ANALYSIS OF THE MARRIAGE TRANSFERENCE MODEL

1. *The Presenting Problem in Interpersonal Terms*

This involves a sample of everyday interpersonal marital conflict, Figure 1 (a), in which the transactions are from husband to wife and wife to

husband, before any differentiation of subselves—one "total" human being interacting emotionally with another "total" human being. This is the usual undifferentiated way in which interpersonal problems are presented by patients (and most others). For example, in group psychotherapy with married couples, Mr. and Mrs. *B* (see protocol) may be asked to relive, by role playing, an argument they had the previous evening. He came home tired and unhappy from an unrewarding day at the office. He wanted to "unwind"—to be fed and to be left alone for a while. She, on the other hand, eagerly awaited his arrival. She needed adult contact. She needed to escape being left alone all day with the children and the chores. Because of their conflict of needs and certain preexisting conditions, a fight ensued which ended with both parties in a self-pitying withdrawing position. Communication was disrupted. They parted company.

2. *The Training Setup*

As illustrated in Figure 1 (b), the first step in the actual training consists essentially of conversion of internal structure to intrapsychic "transactional terms. It is hoped that this will make more "motivational sense" to the observer subselves who will be called upon later in Step 4. In the example cited above both Mr. and Mrs. *B* are independently helped by the therapist to discover the characteristically critical role one of their internalized parental subselves takes with the internalized "hurt child." This leads to an investigation of how the "hurt child" seems to cue the protective (often symbiotic) parent to defend it against the onslaughts of the critical parent. The child may even provoke anger or criticism from the critical parent in order to "get love" from the symbiotic parent who then responds with a lusty shout: "How dare you strike my child!" (Figure 1 (b), thick lines). Although the observer subselves of each spouse are established at this point, they are more or less "just there"—relatively untrained. With both Mr. and Mrs. *B* the transaction is intrapsychic and mostly between opposing parents and each parent with the child. It is interesting to note that the origins of such internal parental conflict can most often be traced to an actual ambivalent parent who played both critical and symbiotic roles in the patient's childhood. Ego therapy helps to make the ambivalent parent discernible by splitting it into two subselves precisely because half of the ambivalence is so often projected onto a spouse and therefore cannot be understood as true ambivalence.

3. *The Problem in Subself Terms*

In the third step husband and wife confront one another, both casts of subselves become interrupted from their internal games and engage externally

(interpersonally) in the game of "How dare you strike my child!" Figure 1 (c) illustrates how Mr. B's critical internal parent begins to turn his fire onto Mrs. B's hurt child, thus temporarily freeing him of his own self-criticism and freeing her of her self-criticism. The *internal* parental-child transactions at this point in both spouses are both protective. From Mrs. B's standpoint Mr. B becomes the villain replacing her own "private" villain—she begins to be freer to feel sorry for herself and really to develop her symbiotic mother-child relationship. Mrs. B's symbiotic mother feels righteously indignant toward her husband and lovingly protective toward her "poor hurt child." He responds in kind with the same cry: "How dare you strike my child!" Again at this point the observers of both parties are relatively inactive, the four parents squabble, but the essential transaction (Figure 1 (c), thick lines) involves a strengthening of the symbiotic parent-child "antiother" reaction. This can be an actual demonstration rather than merely an interpretation of "transference" in the marriage. It also demonstrates the secondary gains of transference.

4. Observer Training

This consists of observation and analysis of modal interpersonal transactions. Although Steps 1 through 3 can be reversed or changed in sequence, Step 4, observation, is always the last step. It is indispensable for the "payoff" in this type of training. The constructive observers of both wife and husband and the total group are elicited (cathected) by asking each of them to step aside, as it were, and have a look at what has been transpiring. They are all asked to comment on what they saw in the relationship as it appeared in the *behavior* during role playing. This is illustrated in Figure 1 (d) in which the observers of Mr. and Mrs. B confer with one another after making observations on their own protective and critical parents, and their child, respectively. At this point there is very little or no transaction between nonobserver sub-selves—and that is the intention of the method. The observational process is aided particularly by the group and the therapist. The aim here is to increase the ability of both parties and the total group to develop more "ego control" by becoming skilled psychological observers and predictors of themselves in action.

E. ILLUSTRATIVE PROTOCOL

The following is an illustrative protocol taken from a group-therapy demonstration¹ of this method. Although actors rather than actual patients

¹ Tape recording of the writer's presentation at the annual meeting of the Los Angeles Group Psychotherapy Society, March 16, 1963.

played the various roles, it represents a fairly accurate sample of the writer's actual technique. The protocol is an illustration of the "marriage transference" model with the specific game of "How dare you strike my child!"

PROTOCOL INTRODUCTION

Dr. Shapiro. I want to try to demonstrate a method of psychotherapy called "ego therapy" and its application to group therapy with married couples. Essentially, the theory is that the ego of each individual functions as though there were a number of subelves, or a cast of characters within it. It is very much the same idea as Eric Berne's ego states—only I call them subelves. Some of these characters are id-like characters; some are ego-like characters in that they carry the classical functions of the ego and others are more superego-like characters, but nevertheless they often exist within us as fairly well-defined characters. I will try to demonstrate that the way in which these subelves interact internally and externally are very important in determining social behavior.

STEP 1. THE INTERPERSONAL PROBLEM (MARITAL CONFLICT)

The first scene is a scene from the home of Mr. and Mrs. Abe Bloom. We see Mrs. Bloom who is at home—just before her husband comes home for dinner. She is preparing dinner as he enters the house.

Mrs. B. Hello, Abe.

Mr. B. Hi . . . (sighs).

Mrs. B. Didya have a hard day?

Mr. B. Yeah . . . (sits down) . . . dinner ready?

Mrs. B. It'll be ready in just a minute.

Mr. B. How come it ain't ready yet?

Mrs. B. (Loud) It'll be ready in just a minute, Abe . . . it's warming up.

Mr. B. When I come home from work, I expect dinner ready.

Mrs. B. It's ready already! I mean I can't wait at the window and wait for you to get out of your car and start walking up to the door.

Mr. B. (Complaining) I get home every day at 6:30. There is no reason supper can't be on the table.

Mrs. B. (Loud) Sometimes you get home at 6:35, sometimes you get home at 6:25—you bawl me out if supper is cold.

Mr. B. What have you got for supper?

Mrs. B. Stew.

Mr. B. Is that the best you could do?

Mrs. B. Yes, it is the best I can do. You don't give me enough money to do anything else. I do the best I can do. You never give me any help.

Mr. B. What kind of help do you want me to give ya, d'ya want me to cook supper for you?

Mrs. B. (Loud, almost tearful) Well, at least you could sit down and help me with the budget sometimes.

Mr. B. I give you eighty bucks a week, that's enough for anybody.

Mrs. B. It's not enough to run a household with children—with three children.

Mr. B. If you bought a little more staples and less crap, you'd have enough to feed everybody.

Mrs. B. You just don't appreciate me.

Mr. B. C'mon, get it on the table, I'm hungry.

Mrs. B. (Yells) Why don't you get your own dinner! (Tears—leaves with muffled sobbing.)

STEP 2. TRAINING IN INTRAPSYCHIC TRANSACTION

The second scene is a meeting in Dr. S's office—a group therapy meeting with four married couples and one male therapist.

Dr. S. Would you come in please folks? (Couples are seated with Dr. S as in therapy.)

Dr. S. (To group) Hi, how are you this evening?

Dr. S. (To Mr. B) You look a little uncomfortable there, Abe.

Mr. B. I had a hard day's work.

Dr. S. Looks like you're all kind of tired out and disgusted or something.

Mr. B. I work hard all day. I take all kinds of abuse. I come home and have to take it from her (Mrs.) too.

Dr. S. Hmmm. (To Mrs. B) Did something happen today?

Mrs. B. We just had another one of our arguments.

Dr. S. Another one of your arguments? You mean you have these all the time?

Mrs. B. Uhuh.

Dr. S. How do you feel about them?

Mrs. B. Well, it was just another one. We just have arguments—there's nothing we can do about it.

Dr. S. You feel kind of helpless as to what to do—you seem resigned to these arguments.

Mrs. B. Yes, they occur all the time.

Dr. S. Uhmm— (To Mr. B) Are you mad at her?

Mr. B. Sure I'm mad. (Sighs) If she kept house . . . the house is dirty when I come home. There's never a friendly word. There's nothing. There's nothing . . . absolutely nothing.

Mrs. B. Well, he comes home and he doesn't talk to me.

Dr. S. Talk to him.

Mrs. B. (To Mr. B) Well, you come home and you don't talk to me, and you don't appreciate me and you bawl me out for not having dinner on the table when you walk in

Mr. B. If I talk to you, you can't carry on a conversation

Mrs. B. . . . the house.

Mr. B. An intelligent conversation. You never could.

Mrs. B. I can! If you'd only

Mr. B. The only thing

Dr. S. Now, let me interrupt this for a moment. (*To Mr. B.*) You said you had a pretty hard day at the office?

Mr. B. Uhuh.

Dr. S. (*To Mrs. B.*) What kind of day did you have before Abe came home? Where were you?

Mrs. B. Well, I was at home, and I ran the errands, you know. Took the children to school—just like every other day.

Dr. S. All right—ok—now, I'd like you (*To Mrs. B.*) to step over to this side, please. Now it's at home. It's in your kitchen. Is that where you were this afternoon before he came home?

Mrs. B. Before he came home I was making dinner.

Dr. S. Oh you were? Would you show us how you were making dinner

Mrs. B. (*Pretends to be stirring pot and working at stove*)

Dr. S. Now, what are you thinking about?

Mrs. B. Just waiting for him to come home and start barking at me.

Dr. S. Do you expect to be barked at?

Mrs. B. Oh yes. He comes home and he's mad. He doesn't want to talk to anybody—just starts barking—we start arguing.

Dr. S. Uhuh. Now, before you start thinking about his barking and his coming home, what is *your* day like? What's your world like before you anticipate his coming home?

Mrs. B. My day . . . my days are just like every day. You know they just Well, I clean the house and I run the children to school and back and

Dr. S. Do you like those things?

Mrs. B. No, no.

Dr. S. Then, why do you do them?

Mrs. B. Because I *have* to. I have a responsibility.

Dr. S. You seem to me like you are kind of trapped in this responsibility. You don't seem very enthusiastic about it.

Mrs. B. Well, I am trapped.

Dr. S. How do you feel about somebody that's trapped? Here sit down (*Mrs. B. sits in one chair directly facing an empty chair*). Let's say that this is Diane. That's your name isn't it? . . . Diane . . . Diane Bloom. Tell her (*the assumed person in the empty chair*) what you think of a person—a woman like this who is trapped.

Mrs. B. (*To assumed self in empty chair*) First of all I think it's your own fault . . . and you should start accomplishing something. You never do anything. I mean you don't even clean the house well and you don't cook well and you could do better and it's *not* all your husband's fault. It's mostly your fault.

Dr. S. What do you think of her? What's your feeling?

Mrs. B. Well, I think she's lazy and I think

Dr. S. Tell her that she's lazy.

Mrs. B. I think you're lazy and I just think that you're not good for anything.

Dr. S. Can you say what you mean by that?

Mrs. B. Well, that she should get out of her rut and start doing something! And start accomplishing something, and really help the children, and help her husband and, uh . . . keep the house clean, and not waste all of her time . . .

Dr. S. All right, now would you sit over here? (*in the empty chair*). Now supposing this were said to you? This (*points to original chair now empty*) is a part of you that's saying this to you. How would you respond to that?

Mrs. B. Well, I do the best I can. It's not . . .

Dr. S. Tell her (*pointing to empty chair*).

Mrs. B. I do the best I can. It's not my fault. It's not all of it my fault. I mean I got married young and no one told me that it was difficult—to be married. I thought it was just . . . you know, you get married . . .

Dr. S. Yes?

Mrs. B. So, all this was plopped onto me all at once and . . . the adjustment starting coming and everything got confused and I . . . everything wasn't all my fault . . . it wasn't all my fault . . . I'm doing the best I can.

Dr. S. You're doing the best you can.

Mrs. B. Yes.

Dr. S. All right, now would you sit over here please? (*a third chair*).

Dr. S. Now, supposing you were an outside observer and you saw these two people talking, what would you see? What can you observe between this person and this person?

Mrs. B. Well, this person . . .

Dr. S. Which one?

Mrs. B. This one (*points to original empty chair*) . . . keeps making excuses and doesn't take any of the criticism and yet this person (*indicates other empty chair*) is criticizing and it's not . . . it's nothing constructive. They're not really communicating.

Dr. S. Now, who is making the excuses for this one? (*points to original empty chair*). She's making her own excuses?

Mrs. B. Uhuh, yes.

Dr. S. Is she sort of defending herself against this criticism?

Mrs. B. Nothing's happening there.

Dr. S. Is she feeling sorry for herself, by the way?

Mrs. B. Yes.

Dr. S. All right, now you be the person right here (*points to a fourth chair*) that feels sorry for this poor girl that's a housewife and has to do all these things. Tell her (*original chair*) how you feel about her.

Mrs. B. Well, you do the best you can, and you try to look nice and you cook with the money you have, and you can't buy steak, and so you make stew and you . . .

Dr. S. What are you saying to her?

Mrs. B. You're doing the best you can.

Dr. S. You're not so bad, huh?

Mrs. B. No-o.

Dr. S. I appreciate you. Is that what you're saying?

Mrs. B. Yes.

Dr. S. Tell her.

Mrs. B. I appreciate you. I think you're doing just fine the way you are.

Dr. S. Ok. Now, would you sit over in this chair please? Now again, as the observer, what do you see here?

Mrs. B. I can see three different personalities.

Dr. S. What are these three different kinds of personalities?

Mrs. B. (*Points to original chair*) Well, I guess this one is the one that's actually moving—actually doing things and making excuses for herself. This one (*points to second chair*) is criticizing, and this one (*points to fourth chair*) is loving . . .

Dr. S. Sort of taking her (*original chair*) part.

Mrs. B. Protecting—Yes, uhuh.

Dr. S. Protecting . . . very good. Thank you very much.

Dr. S. Now, Mr. Bloom, you were talking about a day at the office. Would you sit down there please?

Mr. B. Yeah, uhuh (*sits down*).

Dr. S. Uh, what kind of a day did you have today, Abe?

Mr. B. Oh . . . a day like every other day (*sighs*).

Dr. S. What does that mean?

Mr. B. I get up. I go to work. I look forward for a coffee break, then I'm looking forward for lunch, then I'm looking forward for another break, and then go home . . . nothin' to go home to.

Dr. S. Sounds like there's nothing to go home to, and there's nothing here either.

Mr. B. There's nothing—just nothing.

Dr. S. How does it feel to be in a job and spending your days in nothing . . .

Mr. B. (*Sighs*) It's a useless feeling.

Dr. S. All right, now would you sit over here please? (*second chair*). Here's Abe (*points to first chair*)—right here. Tell him what you think about a guy who spends his days with a useless feeling.

Mr. B. (*To original chair*) Abe, uh, you've got more push. You can do much more than what you're doing. You could utilize your time better. You've got everything that it takes to make good. Both your brothers made good. There's no reason why you shouldn't make good. None at all.

Dr. S. Can you decode that message a little bit—make it a little clearer?

Mr. B. Abe, I admire you. You've got a lot of potential. You do have the push. You show you've got it—so use it.

Dr. S. When you say use it, what feeling are you experiencing?

What is your feeling message?

Mr. B. Find something you can enjoy—something that's worthwhile, when you are not just wasting your time. You're 34 years old, and you should have accomplished something^o already.

Dr. S. Is he a disappointment to you?

Mr. B. Yes, he's a disappointment.

Dr. S. Tell him.

Mr. B. (*Loud*) You're a disappointment to me.

Dr. S. Now, sit over here, please (*original chair*). Supposing you heard that. How would you respond?

Mr. B. It isn't that easy. I've got a wife. I've got three kids. You just can't change a job that easy. If I did change the job how do I know I'd make good on it? There I'd be without a job, and who's gonna feed the kids—take care of the family? There's a lot of bills.

Dr. S. You're saying that he's gotta stay on this job, but you're having a feeling, I suspect, too. What is the feeling?

Mr. B. (*Clears his throat*)

Mr. B. Fear more than anything else. I'm afraid to change jobs, that's all.^o

Dr. S. I think you're saying also, "Don't expect me to change jobs." What do you want from me. I'm afraid. What do you expect from a guy that's afraid.

Mr. B. Yes, that's what I'm saying.

Dr. S. Ok. Now sit here, please (*indicates a third chair*). Now supposing you would be an observer of this . . . what do you see?

Mr. B. Similarity of some sort.

Dr. S. Are they getting along well together, would you say? Are they communicating?

Mr. B. They're communicating, but there's nothing happening.

Dr. S. No changes are taking place.

Mr. B. No changes, no.

Dr. S. How does it make you feel?

Mr. B. Disgust. That's about all.

Dr. S. All right, now tell them both that you're disgusted with them . . . tell them how you feel about them.

Mr. B. You (*to original chair*), you're just using excuses. This is the excuse you used even before you got married and had three kids. It's just an excuse. If it wouldn't be this, it would be something else.

Dr. S. How about this guy? (*points to second chair*).

Mr. B. You . . . What right have you got to criticize like that? None at all.

Dr. S. Uhuh. All right. Now this is, I would say, your observer, the observing part of you that sees this battle going on, and is kind of mad at both of them. Now, would you sit over here please (*indicates original chair*). Is there anything in what he said that was self-pitying? What is this excuse business?

Mr. B. Well, actually . . . I feel . . . sorry for myself, very sorry.

Dr. S. All right sit over here (*indicates a fourth chair*). Now, tell him how sorry you feel for him. Here he is, this poor guy.

Mr. B. Poor guy . . . there you are, you're married to a woman that doesn't understand you . . . that doesn't clean house.

Dr. S. Do your bosses understand you?

Mr. B. Bosses don't understand.

Dr. S. How about your kids? Do they treat you right?

Mr. B. The kids . . . oh . . . what can you expect? She's with them all day long. I barely ever see them.

Dr. S. All right. Now I'm going to talk to some other members of the group. Now we have demonstrated for you, Abe, that there are really four of you. I'm going to ask the other three men to help you by role playing these different parts. Who are you? (*indicates another male group member*).

Male Observer. I'd like to see more. I'm trying to just observe what's going on and control it.

Dr. S. I see. You're the observer and the controller in his life.

Dr. S. (*Points to Mr. B*) You're the person to whom all this is happening. How does it make you feel?

Mr. B. (*Male Child*). It makes me feel very small. Very small.

Dr. S. What makes you feel small about this?

Mr. B. (*Male Child*). Oh, just the idea that I have all these things that I've never discussed with anyone—I have to—just—ah—bare my soul.

Dr. S. It makes you feel about this high? (*indicates height of a three- or four-year-old child*).

Mr. B. (*Male Child*). That's right.

Dr. S. You're the child . . . all right . . . and who are you? (*points to another male group member*).

Male Critic. I'm the critic side of Abe.

Dr. S. Uhuh. What do you say to him usually?

Male Critic. I say to him that he comes to work, and he doesn't look his best. He doesn't have the right attitude.

Dr. S. Uhuh.

Male Critic. When you go to work, you should concentrate on doing your work and forgetting about your problems.

Dr. S. Yes . . . who are you? (*points to fourth male group member*).

Male Protector. I'm more or less the protective parent.

Dr. S. Uhuh, what's your line? What do you do?

Male Protector. I keep him from falling apart completely by telling him that he, ah . . .

Dr. S. Go ahead. Tell him (*Mr. B*).

Male Protector. These other people don't make any difference and ah . . . you are important and you are big and you deserve everything . . . the good life . . . much better life and ah, you shouldn't have to worry about these things. Just take it easy and they'll come.

Dr. S. By the way, what do you think of his wife? Is she treating him right?

Male Protector. Oh no, not at all.

Dr. S. Does he deserve something better than that?

Male Protector. Much, much better . . . she's entirely wrong for him.

Dr. S. Okay, all right, now who are you?

Mrs. B. (Female Child). I'm his wife.

Dr. S. Yes, what's happening to you? You're another person to whom all of this experience is happening. The other side of the marriage . . . the woman. How do you feel when you are in this situation?

Mrs. B. (Female Child). Well, I feel hurt . . . I feel left out . . . I feel that there's . . . I mean he keeps . . . he like . . . he keeps hitting me.

Dr. S. How does that make you feel?

Mrs. B. (Female Child). I feel very small and insignificant.

Dr. S. Insignificant . . . I see . . . like a child.

Mrs. B. (Female Child). Uhuh.

Dr. S. All right, this is the child (*points to Mrs. B.*). Who are you? (*a second female member*).

Female Critic. The critical side.

Dr. S. And what do you have to say?

Female Critic. Well, I just can't stand anything she does.

Dr. S. Well, what about that?

Female Critic. She's lazy and she shouldn't have gotten married this young and I think she's just inferior to her husband . . . and I don't know whether to say disorganized or unorganized . . . anyway, she's so stupid.

Dr. S. She's dumb in addition to it.

Female Critic. Can't even communicate.

Dr. S. And how about you . . . who are you? (*a third female*).

Female Protector. I'm the protective mother, and I disagree because she got married when she was too young, and she lost out on her opportunities, and besides, without her the household would just go to pot; it wouldn't be anything.

Dr. S. All right, and who are you? (*fourth female member*).

Female Observer. I'm the observer. I see these personalities of the four of us and there's an imbalance.

Dr. S. Can you take things into your own hands and really change things? Do you have any control?

Female Observer. If the opportunity is there.

Dr. S. If the opportunity is there.

Female Observer. Except the control is not.

Dr. S. But you don't feel you are able to do much?

Female Observer. No, because nobody sees the other's personality.

STEP 3. HIS SUBSELVES vs. HER SUBSELVES (THE PROBLEM IN SUBSELF TERMS)

Dr. S. Now we have established that there are four characters in each one of these. Now let's come back to the kitchen situation. It is right before the husband is coming home. Would you (*all four women*) come over to the kitchen? The wife is preparing dinner. All right, now what happens?

Female Critic. You know you are lazy . . . that's all you can do is fix stew and you know you can warm it up three nights.

Female Protector. Without you the household would go to pot and wouldn't be anything.

Female Critic. You're just lazy. All you like to do is look pretty for the . . .

Female Protector. You look nice . . .

Female Critic. . . . neighbors downstairs.

Female Protector. You're doing the best you can.

Female Critic. You look terrible in this dress and you don't look the way you used to before you were married.

Female Protector. You could have been a model; you could have been anything; all you would have had to do was try. Why did you get married so young? You gave up everything to marry that pig. If you could be single again you could just jazz around and have fun and not have to have this terrible routine.

Dr. S. Let's stop at this point now. Now it's just before the time when the husband comes home. (*To husband*) You're driving home—you're sitting in your car—another ten minutes and you'll be home. What's happening inside your mind?

Male Protector. It's been a hard day. You're looking forward to a good meal and . . .

Male Critic. You accomplished nothing.

Male Protector. . . . just sit back and relax.

Male Critic. (Sarcastically) You had a drink. Abe, was the beer good, hm? It'll help you now when you get home and have to face your wife, hm? Abe, look at you, you're a mess. No tie. White socks, too. Abe, you forgot your black socks again.

Male Protector. Abe, you have your beer . . . you can relax. Abe, your boss likes you, Abe.

Male Critic. Look, receding hairline, too.

Dr. S. Now let's stop at this moment. These are the states of mind these two people have before they approach one another. Now supposing you (*indicates Mr. B*) take the role of her critic. Would you (*Mrs. B, Female Observer and Female Protector*) (*the Female Critic is out of this*) be in the kitchen again please? Now I'm asking you (*Mr. B*) to role play her critic—remember—*her* critic. You're in the kitchen; he's not home yet.

Mr. B (as her critic). Cooking supper with a \$35.00 dress.

Female Protector. You look nice; you look beautiful; look at those

beautiful shoes. Any man would love to come home to you. You know, you should get out; you shouldn't stay home like that. You can be a model. Show him. Oh, that smells good. You know you are a good cook. Why do you let him boss you around? You're too good for him; you know that. You could have been anything. What is he? What will he ever be?

Mr. B (as her critic). He's too good for you; he's much too good for you.

Dr. S. All right. Let's stop this. What are you (*Mrs. B*) feeling right now?

Mrs. B. Well, I'm being pulled in two different directions. I don't know.

Dr. S. How do you feel toward him? How do you feel toward this guy?

Mrs. B. Very resentful.

Dr. S. Tell him.

Mrs. B. I feel very resentful toward you. You keep shutting me out. You never give me a chance.

Dr. S. (To Mr. B) How do you feel?

Mr. B. I give her as much chance as she asks for. Or do I?

Dr. S. You're not sure what you feel at the present.

Mr. B. At the present time I'm confused.

Dr. S. All right. Now what happens in a marital conflict of this kind is that her critic, who is a very powerful critic; is displaced by him. That is, the critical part of him substitutes, as it were, for the critical part of her. It's like a transference reaction. We've displaced her critical mother and we've inserted him in the role. When he walks into that house and he begins his critical line of her he immediately becomes this critic. She is freed of her critic. The force that her critic internally exerts is no longer there. It's taken by him. So she, really her child, is free then to have sort of a love affair with her protective or symbiotic mother and they can turn their guns on him. Now you are all there and you see this. You (*indicating total group*) are all observers of this. Now, let's replay the original situation—only this time we'll do it with all eight characters. Let's play the original situation when the husband comes home, and the wife is in the kitchen. Okay, would you four (*women*) stand up? You are in the kitchen and you (*men*) come home. Now, remember, that each of the characters has four subelves. Go ahead. You're in the kitchen waiting for your husband.

Female Protector. You poor thing slaving over a hot stove.

Female Critic. Aren't you ashamed of stew again. For goodness sakes, all you can do is look pretty and think about yourself.

Female Protector. You have to put up with so much. You poor thing.

Female Critic. You're really stupid.

Female Child. Oh, here he comes.

Dr. S. Okay. (*All four men come in and sit down as the four women approach them.*)

Mrs. B. (Female Child). Did you have a hard day at the office?

Mr. B. (Male Child). Yeah.

Male Critic. Oh, he did not, he didn't do anything. Three breaks today in twenty minutes.

Mr. B. (Male Child). Is dinner ready?

Mrs. B. (Female Child). It'll be ready in a few minutes.

Mr. B. (Male Child). How come it isn't ready?

Female Critic. Because you're too lazy.

Mrs. B. (Female Child). It has to warm up.

Female Protector. You put up with so much. Don't let him boss you around.

Female Critic. (To Female Protector) Are you kidding?

Male Critic. (Loudly) Look at you, you're a mess. The least you can do is clean up a little bit. You're just a lazy good for nothing!

Female Protector. (Loudly) How dare you strike my child!

Dr. S. What are you going to do with your child that was struck?

Female Protector. Protect her.

Dr. S. Well how?

Female Protector. She's good. We'll all go over in the corner and cry a little for her.

Dr. S. (To Male Protector) Now, what about your child?

Male Critic. Well, look at her. The dinner isn't ready and this house is a mess. I mean what sort of a wife is she?

Male Protector. (Loudly) What are you doing to my child?

Dr. S. What are you going to do now daddy? Your child is hurt.

Male Protector. We're going to console him.

Dr. S. Console him. Go ahead.

Male Protector. Don't worry about it. I mean she's not worth anything. She came to you. She was the one who wanted to get married. You don't. Forget about her. Think about yourself now.

STEP 4. OBSERVER TRAINING

Dr. S. Okay. Now, we're at the last scene of this demonstration and this is kind of in sequence. Now our job is to train the observers of these people. *(To Male Observer)* A lot has happened and you're the observer. You play the observer function in this man's ego, and *(to Female Observer)* you're the observer in this girl's ego. Now, what did you see? What happened, Mr. Observer?

Male Observer. Well, first of all, both individuals are the same in essence and they are rather picture images of each other.

Dr. S. Do you mean mirror images?

Male Observer. Mirror images. Each segment of us has a duplicate in the other individual so therefore we have the same components acting upon us and I as an observer don't have control over us.

Dr. S. All right. Tell them that you don't have control.

Male Observer. I don't have control over you.

Dr. S. How does the observer or person playing this coordinator do his work?

Male Observer. I have to take what seems to me the basic essence of the facts. Like the critic mentioned some points that are true. He may be exaggerating them to a certain extent but there are still some very good points that he gives me. He (*Mr. B*) is rather sloppy; he is consistent in his work sometimes but not all the time. But when I side with the critic in complete bigotry and prejudice against the child, I don't even condescend to see some things that are important. The critic covers up any chance of help. Critic, I would like to be able to point out which are the true points that you have made, but unfortunately I don't have control.

Dr. S. How is it possible for you to get control?

Male Observer. If I had a filter through which all this came; a filter of just an interaction by one individual to what was coming in; if all four of us were toned down to such a degree that there is a level of picking out the major important points.

Dr. S. To sort of keep each person on the same level so that you could pick out what is important.

Male Observer. If the critic swings the lever too far to his side, then, of course, the protective father has to swing it back to his side. But why should this be?

Dr. S. Okay fine. (*To Female Observer*) Now, what did you see in this whole thing? What's the matter with this marriage?

Female Observer. Well, first of all it was brought up before that we are very much alike in the fact that we have these very similar personalities and the fact that the mature individual who is the observer is not functioning.

Dr. S. That means you are not functioning too well among this team, is that it?

Female Observer. I'm watching. I see the imbalance and the other personalities aren't aware of what I am. (*To the others*) "You're just not listening; you're fighting amongst each other. You're not rational."

Female Protector. Well, I can't help it; I just feel stupid. I have to protect her against these horrible . . .

Dr. S. Well, why?

Female Protector. Because otherwise she'd go to pieces. She needs me.

Dr. S. You're saving her. You're important aren't you? (*To Mrs. B*) How do you feel?

Mrs. B. (Female Child). Very confused.

Dr. S. How did you feel about her comment? When she said, "Come on, cut the fighting, you know, let's grow up nice."

Female Protector. I could cry but how can you stop her? (*referring to Female Critic*).

Dr. S. Now this is really a beginning session in group psychotherapy

because at this stage you notice that the observer functions and the coordinating executive functions of their egos are relatively weak—they can't do much. They are mostly observers. Now supposing I put these two observers together and let them be married for a while. You two are married. What are you going to do about this couple—Mr. and Mrs. Observer—what can you do about them?

Male Observer. Well, first of all, I have to realize that we had some faults, meaning us as an individual and that we have to accept these faults and try to compromise with her in some way. Of course, I'd like to see her compromise with us.

Dr. S. Who is the "he" you are talking about?

Male Observer. Her. (*Indicates Mrs. B, the Female Child.*)

Dr. S. What about the observer? What do you want to say to her observer or executive?

Male Observer. Well, I'd like you to try to see some of your faults too and to see some of mine and approach them realistically. We do have to live together, we have to . . .

Dr. S. Why do you have to live together?

Male Observer. Because that's the way our society is and we are married.

Dr. S. Do you want to be married?

Male Observer. Yes, of course we do.

Dr. S. Are you speaking for her?

Male Observer. I am speaking for me, the individual . . .

Dr. S. Do all of you want to be married to her? Do all of the other three?

Male Observer. Yes, in essence they do. These all have to find a release through another individual.

Dr. S. Now what do you (*female observer*) have to say to that?

Female Observer. The fact that you have to recognize that there are these differences now and the fact that I don't have a complete balance and I do have to grow up. The little child is there and it has to be recognized and the only way the little child will grow up into the whole individual is by recognizing this.

Dr. S. (To female observer) Now, I want to ask you a question. Do you think that this team ought to be married to that team?

Female Observer. No.

Dr. S. Why not? Is this a bad marriage in your opinion?

Female Observer. It would be. It would be because they are not a whole human being. They are individual parts.

Dr. S. What about his parts matching your parts? Do they have anything in common? What holds this marriage together?

Female Observer. Well, the fact that they are replacing each other.

Dr. S. You mean it is kind of like a neurotic marriage. Is that what you mean?

Female Observer. It is. It is very neurotic.

Dr. S. Well, he (*indicates male observer*) doesn't see it that way. He feels that there is something good.

Male Observer. I feel there is something solid in that we as a group find release through the different components of her and so therefore it is a good marriage.

Female Observer. But it's not a healthy It's not a healthy marriage.

Dr. S. It's a marriage that sticks together, but that doesn't necessarily mean it's healthy.

F. SUMMARY

The paper extends the ego-therapy approach to the kind of interpersonal transactions found in sensitivity training and the various group therapies. Based on the major theoretical premise of an inner "subself" system in which communication is transactional and markedly similar to interpersonal transaction, this approach also helps to determine at what points such a subself system can react and be changed via interpersonal process. Of the many technique-models which can be derived from this theoretical position, the mirror model, the empathy model, the Eric Berne model and the group as a subself are listed, and the marriage-transference model is illustrated in detail by a protocol.

REFERENCES

1. BERNE, E. *Transactional Analysis in Psychotherapy*. New York: Grove Press, 1961.
2. SHAPIRO, S. B. A theory of ego pathology and ego therapy. *J. of Psychol.*, 1962, **53**, 81-90.
3. ———. Ego therapy in action: A case study. *Psychol. Rep.*, 1962, **11**, 821-831.

Western Psychological Center
4911 Van Nuys Boulevard
Sherman Oaks, California